A Study of Baroclinic Energy Sources for Large-Scale Atmospheric Normal Modes

H. L. TANAKA AND SHAOJIAN SUN*

Geophysical Institute, University of Alaska Fairbanks, Fairbanks, Alaska

(Manuscript received 19 January 1990, in final form 1 June 1990)

ABSTRACT

Observed atmospheric energy peaks in a three-dimensional (3-D) spectral domain are compared with energy peaks predicted by the theory of atmospheric baroclinic instability. The 3-D scale index for global-scale atmospheric motions is represented by the eigenfrequencies of 3-D normal mode functions on a sphere, based on the fact that the eigenfrequencies of Rossby modes are related to the 3-D scale of the waves through the intrinsic wave dispersion relation.

When the observed atmospheric energy level is expressed as a function of the eigenfrequencies, a distinct spectral peak appears in the intermediate value of the eigenfrequencies of Rossby modes. The energy spectrum of atmospheric barotropic components clearly separates a $-5/3$ power law in the high-frequency range, relative to the energy peak, and a 3 power law in the low-frequency range. The peak may describe a certain energy source for large-scale atmospheric motions. For zonal wavenumber 6, we find that the observed spectral peak coincides with the peak predicted from atmospheric baroclinic instability; the energy peak can be produced by baroclinic instability. For zonal wavenumber 2, we also find that the observed spectral peak coincides with that predicted from low-frequency baroclinic instability on a sphere. The results suggest that the low-frequency unstable modes of zonal wavenumber 2 contribute a substantial fraction of the observed spectral peak in a manner similar to zonal wavenumber 6.

1. Introduction

Large-scale atmospheric energy cascades from source ranges to dissipation ranges through wave–wave interactions or turbulence. Kraichnan (1967) and Leith (1968) showed that inertial energy transfers in two-dimensional (2-D) isotropic turbulence allow two equilibrium states: the $-5/3$ and -3 power law inertial subranges. The $-5/3$ range of the two-dimensional flow is known to transfer kinetic energy toward the smaller wavenumbers (larger-scale motions), whereas the -3 range is expected to transfer enstrophy toward the larger wavenumbers (smaller-scale motions). Wiin-Nielsen (1967), using the spectral energetics scheme developed by Saltzman (1957), confirmed the enstrophy-cascading inertial subrange having a -3 slope in the observed atmospheric wind field. There is a maximum baroclinic conversion in synoptic waves, and the kinetic energy is transferred from the source to large- and small-scale motions (Saltzman 1970; Chen and Wiin-Nielsen 1978). The statistical behavior of the 2-D inertial subrange was examined by Lilly (1971), using a numerical experiment of decaying turbulence at a large Reynolds number. His results demonstrated that the initial excitation in the midspectral range tended to exhibit an extensive region having a -3 spectral slope, consistent with Kraichnan's and Leith's results.

Charney (1971) proposed a mathematical isomorphism between the equations governing the 2-D relative vorticity and the 3-D pseudo-potential vorticity. Here, geostrophic turbulence of the three-dimensional quasi-geostrophic flow with a stretched coordinate frame produced a -3 power law. The basic physics of geostrophic turbulence is related to increases in the vorticity of smaller scales due to the elongation of vortex tubes by eddy diffusive separation, which are nonexistent in 2-D turbulence. It is known, however, that Charney's prediction of 3-D isotropization does not hold near the energy source range for synoptic scales. Herring (1980) investigated Charney's geostrophic turbulence in detail, and found that large wavenumbers tended toward three-dimensional isotropy with a log-modified -3 power law spectrum (Kraichnan and Montgomery 1980), while small wavenumbers tended toward a two-dimensional turbulence independent of height (barotropization). Hua and Hiadvogel (1986) also noted discrepancies in Charney's results. Among them a marked difference in the spectral characteristics appeared between barotropic and baroclinic components of the atmospheric turbulence in a variable Brunt–Väisälä stratification. All barotropic modes exhibited a tendency of upscale energy cascade, while all baroclinic modes showed the reverse cascade from large

* Current affiliation: State University of New York at Buffalo.

Corresponding author address: Dr. H. L. Tanaka, Geophysical Institute, University of Alaska—Fairbanks, Fairbanks, AK 99775-0800.

to small scales. Tanaka (1985) and Tanaka and Kung (1988) obtained different spectral slopes for barotropic and baroclinic components of kinetic energy and available potential energy.

Planetary-scale motions have not been considered as fully developed turbulent systems. Theories of turbulence should not be applied to planetary waves (zonal wavenumbers about 1–4), where assumptions of isotropy and homogeneity break down, and triad wave–wave interactions play an important role. For these reasons, theories of 2-D or geostrophic turbulence have been restricted to the smaller scales beyond the source range, zonal wavenumbers 7–20, according to Charney (1971), and have not been extended to planetary waves. Previous research suggests that the spectral characteristics of planetary waves are complicated by many factors, including topographic forcing, land–sea thermal contrast, barotropic and baroclinic instabilities, and nonlinear wave–wave interactions (e.g. Wiin-Nielson 1967; Chen and Wiin-Nielson 1978). Recently, Lilly (1989) discussed the possibility of four inertial subranges for the case of two separate variance sources. These include a large planetary-scale $-5/3$ range, a small synoptic scale -3 range with a downscale enstrophy cascade, a mesoscale $-5/3$ range with an upscale 2-D kinetic energy cascade, and a microscale isotropic $-5/3$ range with a downscale 3-D kinetic energy cascade. This suggests that both the -3 and $-5/3$ power laws of the 2-D inertial subrange exist in planetary-scale motions. The planetary-scale $-5/3$ power law is, however, not confirmed yet from observations.

Using a space–time spectral method, the energetics analysis scheme in the zonal wavenumber domain by Saltzman was extended by Hayashi (1980) to the wavenumber frequency domain. Energy decomposition was further pursued in the meridional direction, using a spherical surface harmonic expansion (e.g. Eliasen and Machenhauer 1965; Baer 1972). Vertically, an expansion in empirical orthogonal functions was employed by several researchers (e.g., Holmström, 1963; Baer, 1981). The three one-dimensional spectral energetics in the domain of zonal wavenumber, meridional index, and vertical index have been combined in the normal mode energetics scheme (e.g. Tanaka 1985; Tanaka and Kung 1988). They applied three-dimensional normal mode functions (3-D NMFs) introduced by Kasahara and Puri (1981) as suitable basis functions for a global energetics analysis. In normal mode energetics, eigenfrequencies of Laplace's tidal equations for various equivalent heights are considered to represent the 3-D scale of the NMFs due to the intrinsic dispersion relation of Rossby waves. According to their results, the energy spectrum of the real atmosphere projected onto these basis functions exhibits a clear, single spectral peak in the eigenfrequency domain. The energy peak separates the region of a 3 power law in the low-frequency range of Rossby modes and the $-5/3$ power law in the high-frequency range of gravity modes. Interestingly, the spectral slope of the largest scale Rossby modes obeys the $-5/3$ law and merges with the gravity modes' $-5/3$ law. The peak is, therefore, not merely associated with the largest scale motions, but it occurs in the intermediate scale. The largest scale $-5/3$ range in the eigenfrequency domain, found by the normal mode energetics, appears to agree with Lilly's prediction by 2-D turbulence.

The interpretation of the distinct energy peak in the intermediate range of the eigenfrequency domain is an open question. It is expected that some energy sources result in the spectral peak. For synoptic disturbances the energy peak may result from atmospheric baroclinic instability, because the growth rates of the unstable modes are large (e.g. Young and Villere 1985; Tanaka and Kung 1989). Although baroclinic instability plays an essential role in synoptic to short waves (zonal wavenumbers about 5–15), the instability has been thought to play a secondary effect in planetary waves (zonal wavenumbers about 1–4) because their growth rate is so small. Hence, for planetary waves there is no clear interpretation of the spectral peak and the 3 and $-5/3$ power law in the eigenfrequency domain. The energy source and wave saturation processes for large-scale motions need to be studied and explained.

The purpose of this study is to evaluate energy and energy-source spectra in the 3-D spectral domain resulting from atmospheric baroclinic instability. The energy spectra of the most unstable modes for realistic zonal basic states are compared with observations for zonal wavenumbers 2, 4, and 6. Baroclinic instability on a sphere is examined, using the 3-D spectral primitive equations with the basis of vertical structure functions and Hough harmonics, as derived by Tanaka and Kung (1989). The stability of the monthly mean, zonal wind field for January 1979 was extensively examined, and compared with the stability of a 30°-jet profile as analyzed by Simmons and Hoskins (1976). The spectral truncation in this study has been extended to include not only the Rossby mode basis as in Tanaka and Kung (1989), but also a comparable number of the gravity mode basis. For comparison, the observed energy spectra based on the analysis of normal mode energetics during the First GARP Global Experiments (FGGE) were reproduced (Tanaka 1985; Tanaka and Kung 1988). It will be shown that the energy source due to baroclinic instability coincides with the observed energy peaks in the eigenfrequency domain both for synoptic and planetary waves.

Section 2 explains the governing equations and a derivation of an algebraic eigenvalue problem to be solved numerically. Section 3 describes the observed energy spectrum in the vertical wavenumber and eigenfrequency domain. The observed energy spectrum with a distinct spectral peak is compared in section 4 with the theoretically expected energy peak due to atmospheric baroclinic instability on a sphere, for both synoptic and planetary waves. Section 5 discusses the

interpretation of the observed spectral peaks that coincide with the theoretically expected energy peaks in planetary waves. Concluding remarks in section 6 summarize the results of this study.

2. Governing equations

A system of primitive equations in spherical coordinates of longitude λ, latitude θ, normalized pressure $\sigma = p/p_b$, and normalized time $\tau = 2\Omega t$ can be reduced to three prognostic equations of horizontal motions and thermodynamics. The three dependent variables are the horizontal wind speeds, $\mathbf{V} = (u, v)$, and a geopotential perturbation, ϕ, from a global reference state. Using a matrix notation, these equations (see Tanaka and Kung 1989) may be written as

$$\mathbf{M}\frac{\partial}{\partial \tau}\mathbf{U} + \mathbf{L}\mathbf{U} = \mathbf{N} + \mathbf{F}, \tag{1}$$

where

$$\mathbf{U} = (u, v, \phi)^{\mathrm{T}}, \tag{2}$$

$$\mathbf{M} = 2\Omega\ \mathrm{diag}\left(1, 1, -\frac{\partial}{\partial\sigma}\frac{\sigma^2}{R\gamma}\frac{\partial}{\partial\sigma}\right), \tag{3}$$

$$\mathbf{L} = \begin{pmatrix} 0 & -2\Omega\sin\theta & \dfrac{1}{a\cos\theta}\dfrac{\partial}{\partial\lambda} \\ 2\Omega\sin\theta & 0 & \dfrac{1}{a}\dfrac{\partial}{\partial\theta} \\ \dfrac{1}{a\cos\theta}\dfrac{\partial}{\partial\lambda} & \dfrac{1}{a\cos\theta}\dfrac{\partial(\ )\cos\theta}{\partial\theta} & 0 \end{pmatrix}, \tag{4}$$

$$\mathbf{N} = \begin{pmatrix} -\mathbf{V}\cdot\nabla u - \omega\dfrac{\partial u}{\partial\sigma} + \dfrac{\tan\theta}{a}uv \\ -\mathbf{V}\cdot\nabla v - \omega\dfrac{\partial v}{\partial\sigma} - \dfrac{\tan\theta}{a}uu \\ \dfrac{\partial}{\partial\sigma}\left(\dfrac{\sigma^2}{R\gamma}\mathbf{V}\cdot\nabla\dfrac{\partial\phi}{\partial\sigma}\right) \end{pmatrix}, \tag{5}$$

and

$$\mathbf{F} = \left(F_u, F_v, \frac{\partial}{\partial\sigma}\frac{\sigma Q}{c_p\gamma}\right)^{\mathrm{T}}. \tag{6}$$

The symbols in these equations are the earth's radius a, the angular speed of the earth's rotation Ω, the specific gas constant R, and the specific heat at a constant pressure c_p. Also defined are the static stability parameter γ, the normalized vertical p-velocity $\omega = d\sigma/dt$, the zonal and meridional components of the frictional force F_u, F_v, and the diabatic heating rate Q. Based on scale considerations involving ω, the vertical advection of the temperature perturbation has been neglected in (5) as in Kasahara and Tanaka (1989).

The 3-D NMFs are obtained as eigensolutions of the generalized eigenvalue problem given by (1) with $\mathbf{N} + \mathbf{F} = 0$, subject to prescribed boundary conditions. Periodicity along longitudinal circles and a regularity of the solution at the north and south poles are the usual boundary conditions in the horizontal. For geometrical simplicity, the atmosphere is treated as a unit sphere with the radius 1 defined over a domain of $0 < \sigma \leqslant 1$. The upper boundary condition is given at the neighborhood of the origin of the unit sphere, and the lower boundary condition at the surface of $\sigma = 1$. Using the isobaric coordinate formulated by Boer (1989), the domain of the equation is extended to $0 < p \leqslant p_b$, where p_b is a *fixed* value greater than the surface pressure p_s at all times.

The lower boundary condition for the atmosphere is derived from the condition of a vanishing wind vector that consists of horizontal u, v, and vertical $w = dz/dt$ wind speeds:

$$(u, v, w) = 0, \quad \text{at} \quad \sigma = 1. \tag{7}$$

This condition always holds within the "subterranean" regions for which $p_s < p < p_b$. Since three dependent variables in (2) constitute a vector, three boundary conditions for the general nonlinear system must be considered as in (7). We assume that the diabatic heating also vanishes within the subterranean region so that temperature is extrapolated adiabatically, as is commonly practiced. The subterranean values have little effect on energy calculations, because the key variables are multiplied by the Heaviside or unit function such as

$$\beta(p - p_s) = \begin{cases} 1, & \text{if} \quad p < p_s; \\ 0, & \text{if} \quad p > p_s. \end{cases} \tag{8}$$

The usual operations of calculus, extended to the behavior for this unit function and its derivative (Dirac delta function, $\partial\beta/\partial p = -\delta$) permits further derivations of the model atmosphere (see Boer 1989). In this formulation, ω at $p = p_b$ is identical to the value at $p = p_s$ and has a nonzero value. These conditions, combined with the first law of thermodynamics, lead to the lower boundary condition for the vertical structure function $G(\sigma)$:

$$\frac{dG}{d\sigma} + \frac{\gamma_b}{T_b}G = 0 \quad \text{at} \quad \sigma = 1, \tag{9}$$

where the subscript b denotes the value of the reference state at $\sigma = 1$.

The upper boundary condition is given by statements that bound total kinetic and available potential energies:

$$\int_0^1 K + A\,d\sigma = \int_0^1 \frac{1}{2}(u^2 + v^2) + \frac{\sigma^2}{2R\gamma}\left(\frac{\partial\phi}{\partial\sigma}\right)^2 d\sigma < \infty. \tag{10}$$

These conditions lead to the upper boundary condition for the vertical structure function $G(\sigma)$:

$$\int_0^1 G^2 d\sigma < \infty. \tag{11}$$

We have assumed that γ is positive. For the case of a constant static stability, the vertical structure equation reduces to an Euler equation, as seen in the radial equation of spherical harmonics. An isothermal atmosphere is a special case of a positive constant static stability. The solutions are power functions with a real negative power for a barotropic (external) solution and complex powers for baroclinic (internal) solutions. For an open atmosphere with a singularity at $\sigma = 0$, rigorous mathematical treatment of the singular Sturm–Liouville boundary value problem yields a continuous spectrum for internal solutions and a discrete spectrum for the external solution (see Staniforth et al. 1985). However, the continuous spectrum causes difficulties in further development. In this study, the solutions are constructed numerically, according to Galerkin's procedure as used by Kasahara (1984) to obtain bounded vertical structure functions with a discrete spectrum. In his scheme, the vertical structure functions are approximated by a finite series of Legendre functions. The vertical structure equation is then solved as an algebraic matrix eigenvalue problem. The boundedness, together with the lower boundary condition given by (9), is sufficient for the so-called mass matrix operator in (3) to be selfadjoint. The accuracy of the numerical solution thus depends on the matrix size. The finite number of eigenvalues are regarded as a discrete approximation of the continuous spectrum.

An equivalent height, h_m, which appears as a separation constant for the vertical and horizontal structure equations, is proportional to the inverse of the eigenvalue of the vertical structure equation. It describes the scale of vertical normal modes as the spectrum forms a monotonically increasing sequence of eigenvalues. A horizontal structure equation (that is, Laplace's tidal equation) is then solved for the given separation constant of the equivalent height to obtain Hough harmonics $H_{nlm}(\lambda, \theta)$. The 3-D NMFs, $\Pi_{nlm}(\lambda, \theta, \sigma)$, are defined as a tensor product of the vertical structure functions $G_m(\sigma)$ (vertical normal modes) and Hough harmonics $H_{nlm}(\lambda, \theta)$ (horizontal normal modes)

$$\Pi_{nlm}(\lambda, \theta, \sigma) = G_m(\sigma)H_{nlm}(\lambda, \theta) = G_m(\sigma)\Theta_{nlm}(\theta)e^{in\lambda}. \tag{12}$$

Here, $\Theta_{nlm}(\theta)$ denotes Hough vector functions, and the subscripts represent zonal wavenumbers n, meridional indices l, and vertical indices m (refer to Kasahara and Puri 1981; Tanaka 1985, for greater detail). These can be partitioned in rotational (Rossby modes l_R) and gravitational modes (eastward l_E and westward l_W gravity modes) based on the eigenfrequency, σ_{nlm}, of the generalized eigenvalue problem. One must be careful to avoid confusion when σ_{nlm} is used for eigenfrequency and σ for the vertical coordinate.

The generalized eigenvalue problem derived from the matrix operators **M** and **L** in (1) can be written in a dimensionless form as

$$i\sigma_{nlm}(\mathbf{Y}_m^{-1} \quad \mathbf{M} \quad \mathbf{X}_m)\Pi_{nlm} = (\mathbf{Y}_m^{-1} \quad \mathbf{L} \quad \mathbf{X}_m)\Pi_{nlm}. \tag{13}$$

The scaling matrices, $\mathbf{X}_m$ and $\mathbf{Y}_m$ are defined using $c_m = \sqrt{gh_m}$, where g is the earth's gravity, for each vertical index:

$$\mathbf{X}_m = c_m \,\mathrm{diag}(1, 1, c_m), \tag{14}$$

$$\mathbf{Y}_m = 2\Omega \,\mathrm{diag}(c_m, c_m, 1). \tag{15}$$

The geometry and stratification of the real atmosphere confront us with the problem of defining a 3-D scale. One approach to solving this problem is to derive a differential equation that describes the atmosphere to some extent. The 3-D scaling can be established based on the ordered eigenvalues of the differential equation (see Baer 1981). Equation (13) has a standard form of the generalized eigenvalue problem of a differential equation. Hence, the ordered eigenfrequency can be regarded as a three-dimensional scale index of the 3-D NMF. The eigenfrequency of a Rossby mode is related to its 3-D scale through the wave dispersion relation. The magnitude of the eigenfrequency decreases, in general, as n, l, and m increase; that is, as the 3-D scale decreases. Although there is an exception where the relation reverses at very small n, the eigenfrequency may be understood as a 3-D scale index. On the other hand, the eigenfrequency of a gravity mode decreases in magnitude as m increases, but it increases as n and l increase. Therefore, when different vertical modes are compared, the 3-D scaling for a gravity mode is not as simple as a Rossby mode.

It should be noted that the zonal wavenumber n is a square root of an eigenvalue of a longitudinal differential equation, which describes the longitudinal scale. It is also an integer that counts the sequence of the eigenvalues. In contrast, the eigenvalue of the vertical differential equation is $(h_m)^{-1}$ (apart from a constant g), and the vertical index m is an arbitrary integer that counts the sequence of the eigenvalues. It may start from $m = 1$ instead of $m = 0$ for the barotropic component. That is, the eigenvalue no longer coincides with the index. Taking account of the distinction between the invariant $(h_m)^{-1}$ and the free index m, we propose to use $(h_m)^{-1}$ for the quantitative measure of the vertical scale, and designate this a vertical wavenumber. Similar arguments can be applied to the eigenfrequency σ_{nlm} and the meridional index l. When both n and m are fixed, the eigenfrequency represents a one-dimensional meridional scale. The meridional index l of Hough harmonics is an arbitrary index that may start from other than $l = 0$. In this study, we will use the

eigenfrequency σ_{nlm} as a quantitative measure of the meridional scale.

It can be shown by the Sturm–Liouville theorem that the 3-D NMFs are mutually orthonormal with a natural inner product:

$$\langle \Pi_{nlm}, \Pi_{n'l'm'} \rangle = \delta_{nn'}\delta_{ll'}\delta_{mm'}, \quad (16)$$

and that these form a complete set of expansion basis functions for arbitrary variables. Based on this property, we approximate the variables **U** and **F** in (1) by a finite series of the 3-D NMFs:

$$\mathbf{U}(\lambda, \theta, \sigma, \tau) = \sum_{n,l,m} w_{nlm}(\tau)\mathbf{X}_m\Pi_{nlm}(\lambda, \theta, \sigma), \quad (17)$$

$$\mathbf{F}(\lambda, \theta, \sigma, \tau) = \sum_{n,l,m} f_{nlm}(\tau)\mathbf{Y}_m\Pi_{nlm}(\lambda, \theta, \sigma). \quad (18)$$

Here, the expansion coefficients w_{nlm} and f_{nlm} are functions of only time. Using the orthonormality condition of (16), the expansion coefficients can be obtained from

$$w_{nlm} = \langle \mathbf{U}, \mathbf{X}_m^{-1}\Pi_{nlm} \rangle, \quad (19)$$

$$f_{nlm} = \langle \mathbf{F}, \mathbf{Y}_m^{-1}\Pi_{nlm} \rangle. \quad (20)$$

It is important to note that the expanded variables, including the zonal field, are in the subspace spanned by the 3-D NMFs. Therefore, they satisfy the same boundary conditions of the 3-D NMFs of (9) and (11), which were derived from (7) and (10). It is found from (9), that if $V = 0$, then $\partial V/\partial\sigma = 0$ at $\sigma = 1$. Hence, this condition permits the internal baroclinic instability as is discussed by Lindzen et al. (1980) in the concept of the over reflection (see Tanaka and Kung 1989; Kasahara and Tanaka 1989). Since the boundary condition given by (9) forbids the boundary effect in the Charney–Eady problem, the instability in this study appears to occur by a negative meridional gradient of the potential vorticity of the basic state near the ground (see the Charney–Stern theorem; Holton 1975).

Applying the inner product for (1) and using the scale matrices in (14) and (15), we can construct a weak form of the primitive equation:

$$\left\langle \mathbf{M}\frac{\partial}{\partial\tau}\mathbf{U} + \mathbf{L}\mathbf{U} - \mathbf{N} - \mathbf{F}, \mathbf{Y}_m^{-1}\Pi_{nlm} \right\rangle = 0. \quad (21)$$

From this, the 3-D spectral representation of the primitive equation becomes a system of ordinary differential equations for the expansion coefficients given in (17):

$$\frac{dw_i}{d\tau} + i\sigma_i w_i = -i\sum_{j=1}^{K}\sum_{k=1}^{K} r_{ijk}w_j w_k + f_i, \quad i = 1, 2, \cdots, K, \quad (22)$$

where K is the total number of the series expansion. For simplicity, the three triple subscripts nlm, $n'l'm'$, $n''l''m''$ have been shortened to subscripts i, j, and k, respectively. The real interaction coefficients r_{ijk} are explicitly evaluated by the triple product of the basis functions, according to (21).

In order to analyze the stability of a zonal basic state, the nonlinear equation (22) is linearized for the prescribed basic state, using notations $\bar{w}_i$ and $\bar{f}_i$ for the time-independent zonal basic state and w_i and f_i for small perturbations superimposed on the basic state (the same symbols as the original variables are used for convenience). The real physical implication of the linearization will be reconsidered in section 5. The resulting system of linear equations for the first-order terms of the perturbation quantities can be expressed in a matrix form. For inviscid and adiabatic perturbations, it becomes

$$\frac{d}{d\tau}\mathbf{W} + i\mathbf{D}\mathbf{W} = -i\mathbf{B}\mathbf{W} \quad (23)$$

where

$$\mathbf{W} = (w_1, w_2, \cdots, w_i, \cdots, w_N)^{\mathrm{T}}, \quad (24)$$

$$\mathbf{D} = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_i, \cdots, \sigma_N). \quad (25)$$

The dimension of the system **N** is the product of the total vertical indices and total meridional indices. The (i, j) entry of the real matrix **B** is determined by the expansion coefficients of the zonal basic state:

$$b_{ij} = \sum_{k=1}^{N}(r_{ijk} + r_{ikj})\bar{w}_k, \quad i, j = 1, 2, \cdots, N. \quad (26)$$

If we assume that the solution of **W** is

$$\mathbf{W}(\tau) = \xi\exp(-i\nu\tau), \quad (27)$$

the initial value problem (23) becomes an eigenvalue problem for a real matrix $(\mathbf{D} + \mathbf{B})$ with the eigenvector ξ and eigenvalue ν as

$$\nu\xi = (\mathbf{D} + \mathbf{B})\xi. \quad (28)$$

The real and imaginary parts of the eigenvalue ν give the frequency and growth rate. By solving (28) for the prescribed zonal basic state, we can examine the unstable modes that provide the energy source for synoptic and planetary waves.

Finally, according to Kasahara and Puri (1981) and Tanaka (1985), an energy element E_i for a particular basis function is defined in a dimensional form by

$$E_i = \frac{1}{2}p_b h_m |w_i|^2. \quad (29)$$

The energy spectra of the observed atmosphere and the analyzed unstable modes can be examined using (29). It can be shown that the energy-source spectrum of the unstable modes, dE_i/dt, is proportional to the energy spectrum itself. Using ν_I as the positive imaginary part of the eigenvalue, it becomes

$$\frac{dE_i}{dt} = 4\Omega\nu_I E_i. \quad (30)$$

Since the expansion basis is composed of Rossby modes and gravity modes, we can describe the spectral distributions of the energy and energy source for the rotational and gravitational components as functions of the eigenfrequency, $|\sigma_i|$. The energy and energy source can also be expressed as functions of the vertical wavenumber, h_m^{-1}.

3. Large-scale atmospheric energy spectrum

a. Energy spectrum in the zonal wavenumber domain

Atmospheric energy levels are generally higher for planetary waves (n = 1–4) than for synoptic to short waves (n = 5–15) (e.g., Saltzman 1970; Kung 1988). Figure 1 compares the spectral distributions of total energy for the transient component of motions, E_T, and the sum of the stationary component and annual cycle, $E_S + E_A$ (after Tanaka and Kung 1988). Here, the total energy is the sum of the kinetic energy and available potential energy. The results of the normal mode energetics (solid lines) are compared with the standard spectral energies (dashed lines), based on the one-year period of the FGGE III-b data assimilated at the Geophysical Fluid Dynamics Laboratory (GFDL). The energy level of E_T is one order higher than that of $E_S + E_A$ for $n > 4$, indicating the dominance of transient waves of synoptic disturbances. The energy level of $E_S + E_A$ becomes comparable to E_T at $n = 1$. The major part of quasi-stationary flow is contained within the range of $n \leq 4$. The results of the transient energy spectrum E_T show a flat energy spectrum for $1 < n < 6$ and a clear transition to the -3 power-law range for $n > 6$. The edge of the -3 power law range near $n = 6$ corresponds to the range of the energy source of synoptic disturbances. Atmospheric baroclinic instability has been considered as the major energy source of the synoptic disturbances. The maximum baroclinic conversion occurs in synoptic transient waves, while the kinetic energy is transformed from the source to large- and small-scale motions (see Saltzman 1970). Recent studies using the FGGE data show an additional maximum of baroclinic conversion in planetary waves associated with low-frequency variability in the Northern Hemisphere (Kung and Tanaka 1984; Kung 1988).

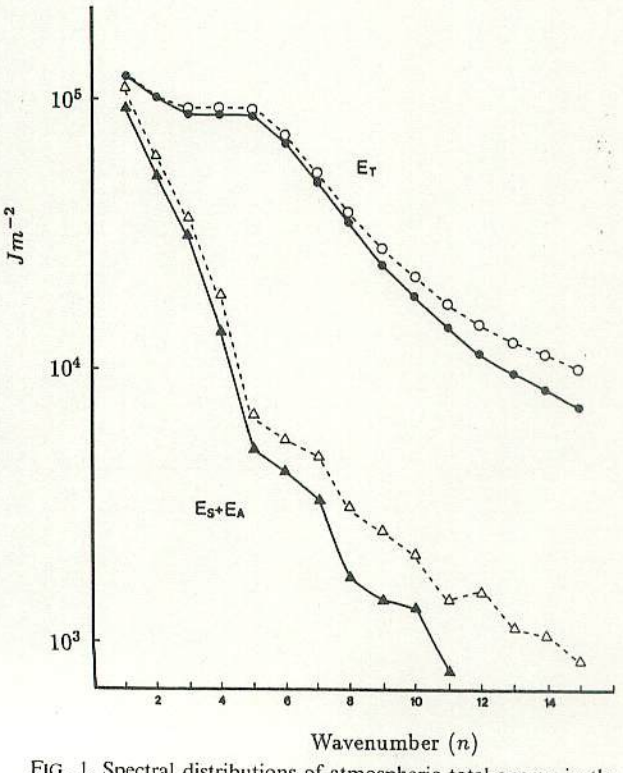

FIG. 1. Spectral distributions of atmospheric total energy in the zonal wavenumber domain for contributions from the steady component plus annual cycle of motions, $E_S + E_A$, and transient motions, E_T (after Tanaka and Kung 1988). The results are based on FGGE observations. The normal mode energetics are solid lines and the standard spectral energetics are dashed lines.

b. Energy spectrum in the vertical wavenumber domain

The vertical wavenumber is proportional to Lamb's parameter ($=a^2\Omega^2/gh_m$). However, it is not directly proportional to the vertical geometrical wavenumber. Silva Dias and Bonatti (1986) demonstrated that the vertical energy spectra, evaluated with various sets of vertical structure functions with different upper boundary conditions, are approximately unique over $(h_m)^{-1}$. Atmospheric energy spectra in the vertical wavenumber domain are illustrated in Fig. 2 for zonal wavenumbers n = 0, 2, 4 and 6. The evaluation of total energy is based on the FGGE winter data (December, January, and February), using (2), (19), and (29). For all meridional indices, the energy elements E_i are summed separately for Rossby modes, eastward gravity modes, and westward gravity modes. The resulting energy levels are plotted as functions of the vertical wavenumber, h_m^{-1}, in a unit of m^{-1}. Table 1 lists the correspondence between the vertical indices and the vertical wavenumbers.

The rotational part of the zonal component n = 0, is called the geostrophic mode (Kasahara 1978). The highest energy (23.4×10^5 J m^{-2}) is seen at $h_m^{-1} = 7 \times 10^{-3}$. Since the total energy for n = 0 is about 60×10^5 J m^{-2} (Kung and Tanaka 1983), close to half the energy is stored near the peak at $h_m^{-1} = 7 \times 10^{-3}$. The barotropic mode at $h_m^{-1} = 1 \times 10^{-4}$ contains 8.4×10^5 J m^{-2}. Another energy peak is seen at $h_m^{-1} = 5 \times 10^{-2}$. These are energy gaps between $h_m^{-1} = 1 \times 10^{-4}$ and 7×10^{-3} and between 7×10^{-3} and 5×10^{-2}. The gravity mode energy is two orders of magnitude less than the energy of the geostrophic mode. Two energy peaks appear in the vertical wavenumbers as seen in the geostrophic modes. Since both of v and ω are identically zero in the geostrophic modes, important information concerning the Hadley circulation is contained in the gravity modes although the energy levels are low.

For wavenumbers n = 2, 4, and 6, the Rossby modes indicate energy peaks at $h_m^{-1} = 1 \times 10^{-4}$ and 7×10^{-3} with a weak peak at $h_m^{-1} = 5 \times 10^{-2}$. The energy

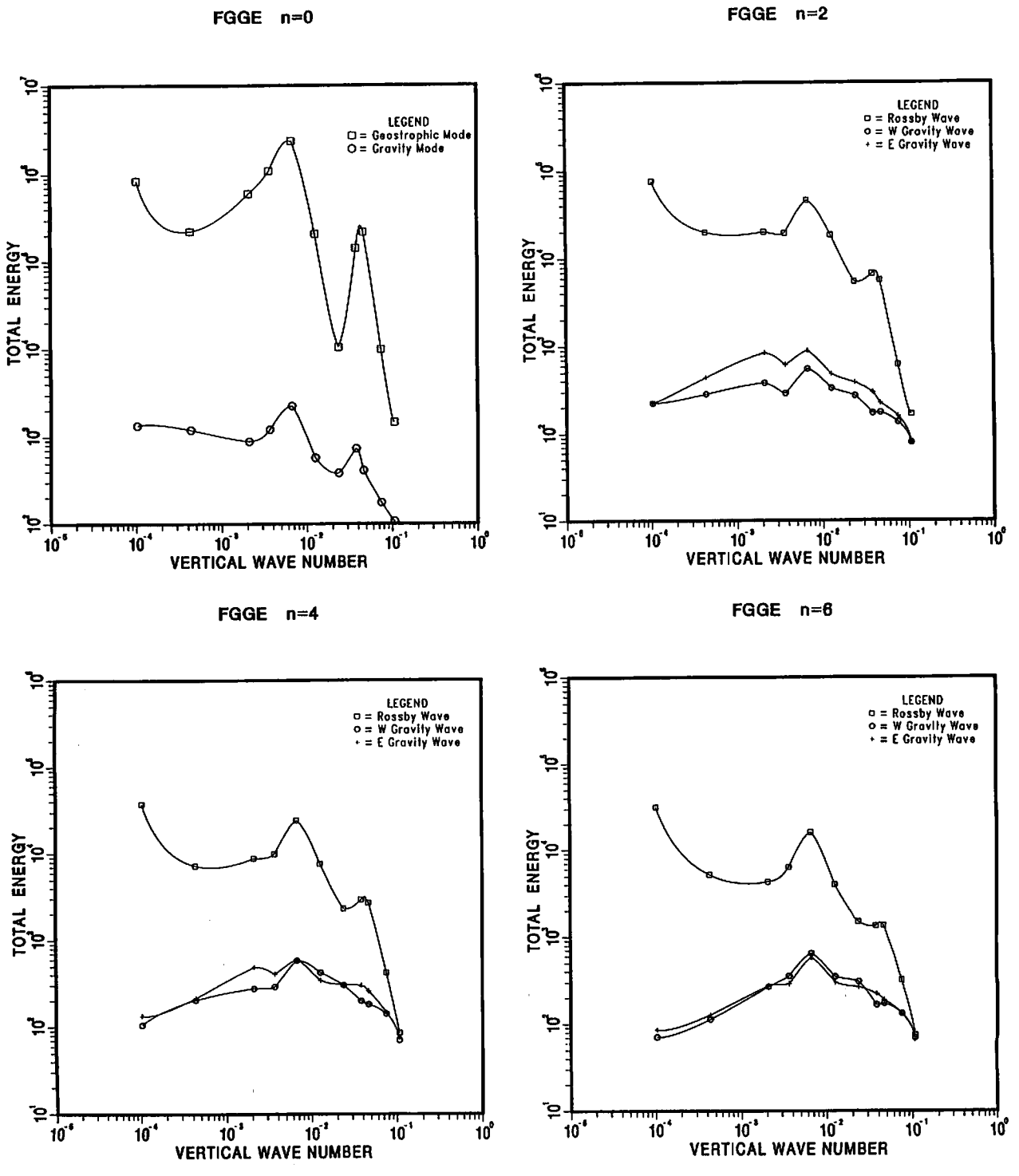

FIG. 2. Spectral distributions of atmospheric total energy in the vertical wavenumber domain for zonal wavenumbers n = 0, 2, 4 and 6 during the FGGE winter. The vertical wavenumber h_m^{-1} (unit: m^{-1}) is the inverse of the equivalent height, with m denoting the vertical indices 0 through 10. The symbols designate Rossby modes (squares), westward gravity modes (circles), and eastward gravity modes (crosses). For n = 0, the sum of positive and negative frequency modes is designated as gravity modes.

levels decrease rapidly in the higher vertical modes beyond $h_m^{-1} = 7 \times 10^{-3}$. It is interesting to note that overall, the energy distributions for $h_m^{-1} > 7 \times 10^{-3}$ are very similar for all zonal waves. The energy gaps between $h_m^{-1} = 1 \times 10^{-4}$ and 7×10^{-3} are related with the first internal mode, m = 1, which has its largest amplitude at the model's top. These gaps tend to be filled in by planetary waves. This tendency may be related to the occasional vertical propagation of planetary waves into the deep atmosphere of the strato-

TABLE 1. The vertical index m, the corresponding equivalent height h_m (m), and the vertical wavenumber h_m^{-1} used for the normal mode energetics (left) and atmospheric stability analysis (right) in this study.

m	h_m	h_m^{-1}	m	h_m	h_m^{-1}
0	9623.9	1.0×10^{-4}	0	9746.5	1.0×10^{-4}
1	2297.1	4.4×10^{-4}	1	3457.7	2.9×10^{-4}
2	475.9	2.1×10^{-3}	2	777.5	1.3×10^{-3}
3	272.0	3.7×10^{-3}	3	269.3	3.7×10^{-3}
4	150.0	6.7×10^{-3}	4	113.9	8.8×10^{-3}
5	79.5	1.3×10^{-2}	5	46.9	2.1×10^{-2}
6	42.4	2.4×10^{-2}	6	24.1	4.1×10^{-2}
7	26.3	3.8×10^{-2}			
8	21.6	4.6×10^{-2}			
9	13.4	7.5×10^{-2}			
10	9.4	1.1×10^{-1}			

sphere and mesosphere. On the other hand, gravity mode energy distributions indicate peaks at $h_m^{-1} = 7 \times 10^{-3}$. The energy levels are two orders of magnitude lower than that of Rossby modes for the barotropic component, whereas these become comparable near $h_m^{-1} = 1 \times 10^{-1}$. There is a significant difference in the energy levels of westward and eastward gravity modes for $n = 2$ and 4. The eastward modes contain more energy than the westward modes in the range of $10^{-3} < h_m^{-1} < 10^{-2}$. The difference comes from active tropical Kelvin waves (see Tanaka and Kung 1988).

c. *Energy spectrum in the eigenfrequency domain*

Before introducing the energy spectrum in the eigenfrequency domain, we consider Fig. 3, which shows the energy spectra of the wavenumber $n = 2$ as functions of the meridional indices l_R, l_W, and l_E (see Tanaka 1985, for the detail). As expected, the Rossby modes dominate the gravity modes in the three panels of $m = 0$, 2, and 4. The energy peaks appear at the intermediate meridional indices of $l_R = 3$ for $m = 0$, $l_R = 5$ for $m = 2$, and $l_R = 6$ for $m = 4$. The energy levels of the largest-scale Rossby mode, $l_R = 0$ are very low and are comparable to the energy levels of the gravity modes.

The eigenfrequency σ_{nlm} given in (13) describes the meridional scale of the mode within the same type of Rossby and gravity modes for the fixed n and m. The larger the meridional index, the smaller the model meridional scale (see Kasahara 1976, for the structures). As the meridional index increases, the magnitude of the eigenfrequency $|\sigma_{nlm}|$ decreases monotonically within Rossby modes; whereas it increases monotonically within gravity modes. Consequently, the discrete spectrum of $|\sigma_{nlm}|$ spans the entire positive frequency domain. Mixed Rossby–gravity modes ($l_R = 0$) and Kelvin modes ($l_E = 0$) are positioned at the boundary of the Rossby and gravity modes. The spectral characteristics, such as the power law in the wavenumber

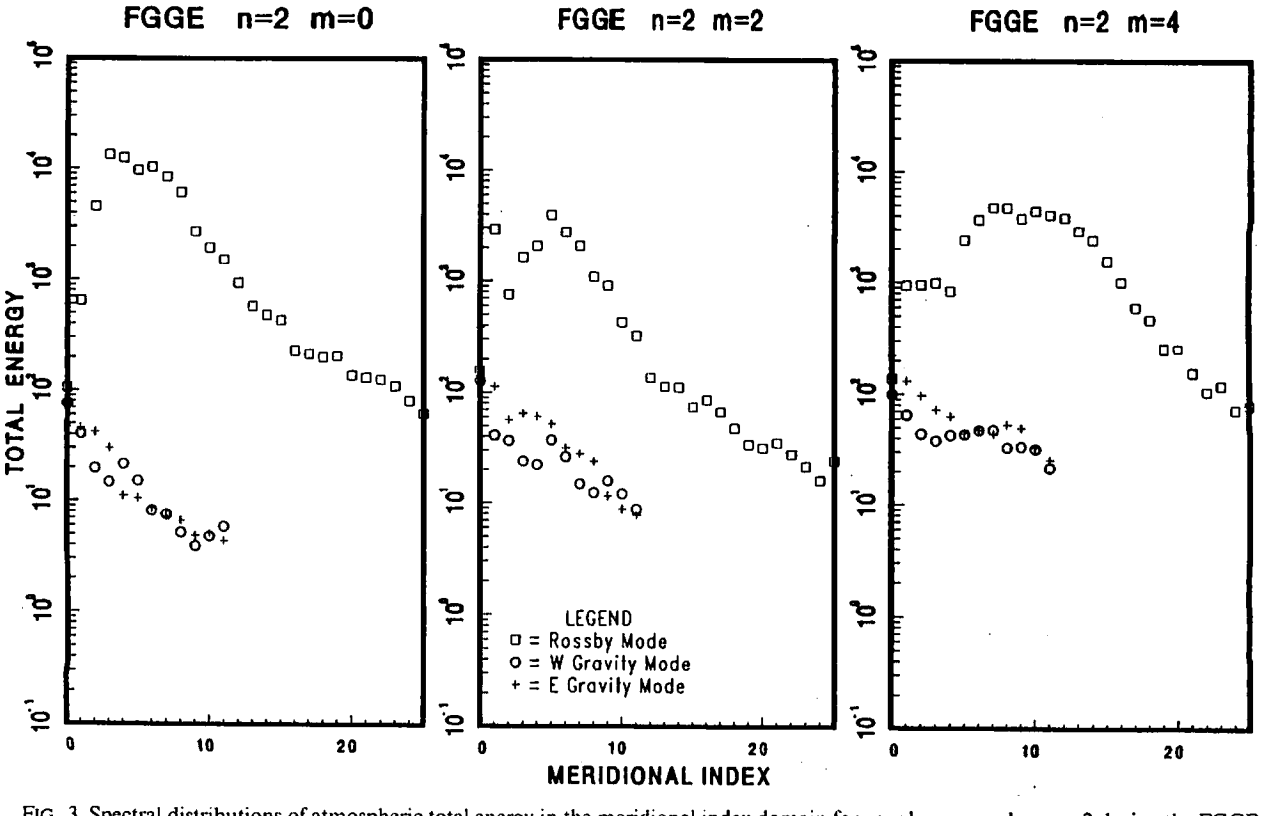

FIG. 3. Spectral distributions of atmospheric total energy in the meridional index domain for zonal wavenumber $n = 2$ during the FGGE winter. The three panels represent for the vertical indices $m = 0$, 2, and 4. Rossby and gravity modes are plotted with different symbols shown in the legend.

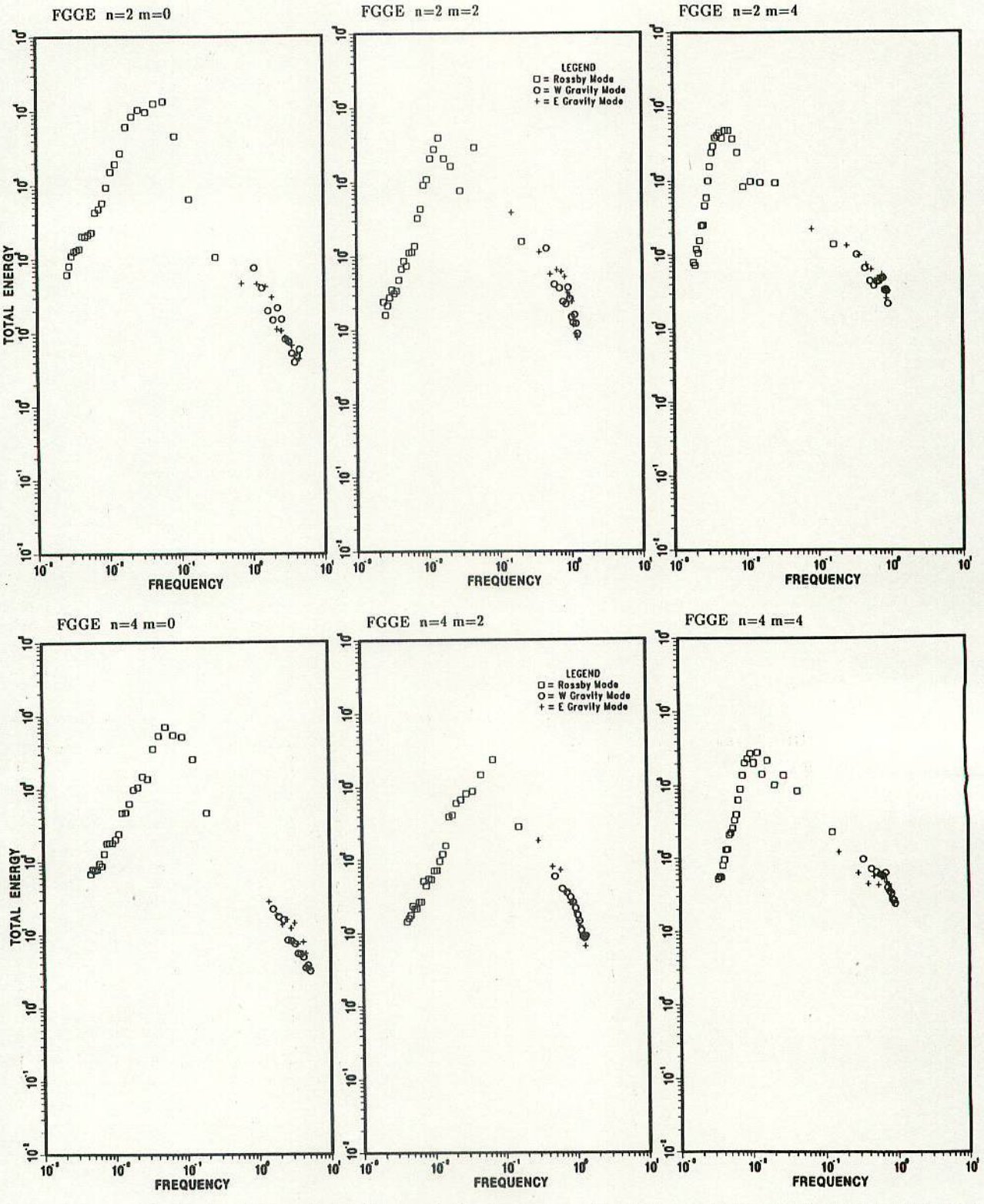

FIG. 4. Spectral distributions of atmospheric total energy in the eigenfrequency domain for zonal wavenumbers n = 2, 4, and 6 during the FGGE winter. The dimensionless eigenfrequencies $|\sigma_{nlm}|$ in (13) are chosen in the abscissa to represent the meridional scale of atmospheric

domain, are transformed to the frequency domain using the wave dispersion relations.

The results of atmospheric energy spectra as functions of $|\sigma_{nlm}|$ are illustrated in Fig. 4 for n = 2, 4, and 6. It should be noted that the eigenfrequency in the abscissa differs from the wave frequency in space–time spectra as in Hayashi (1980). The three panels in each figure describe the energy spectra of the vertical indices

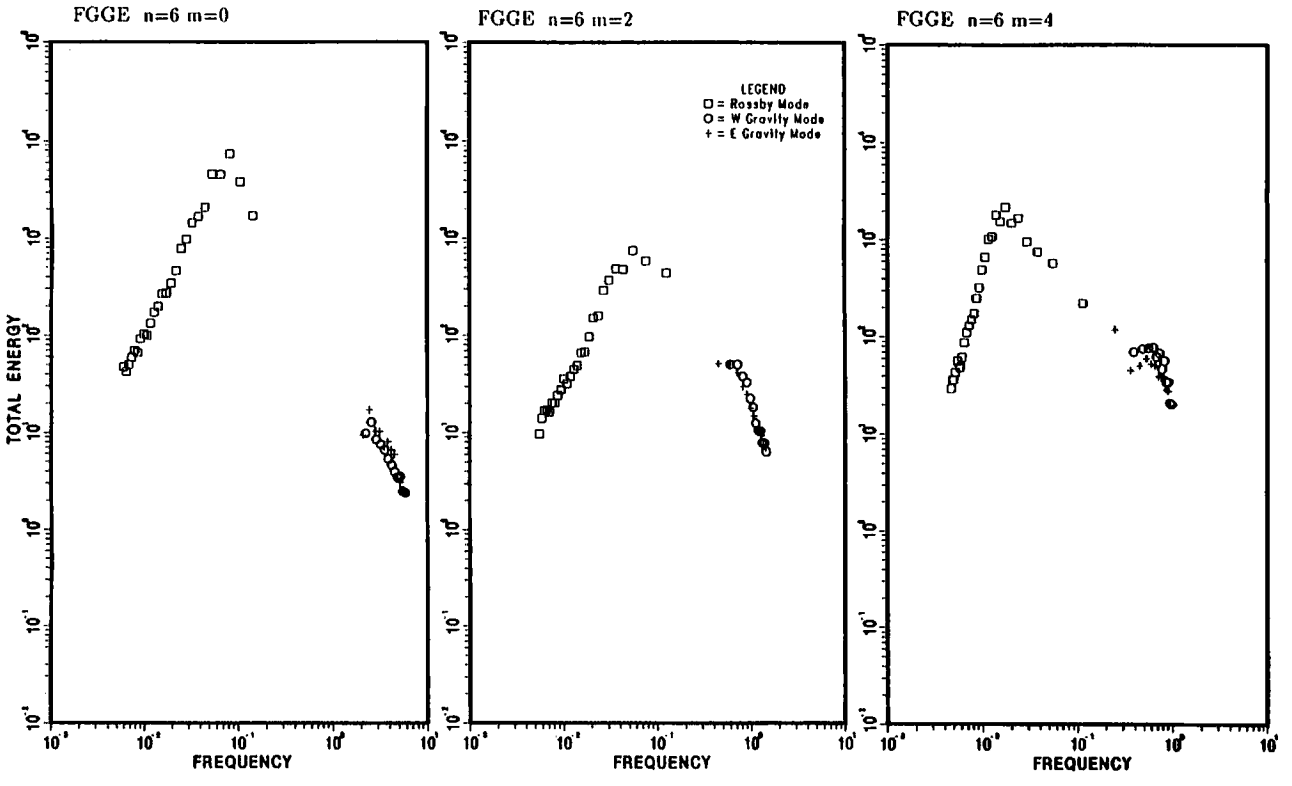

FIG. 4. (*Continued*) motions. The three panels in each figure represent for the vertical indices $m = 0$, 2, and 4. Rossby and gravity modes are plotted in the low-frequency and high-frequency ranges, respectively.

of $m = 0$, 2, and 4, corresponding to $h_m^{-1} = 1 \times 10^{-4}$, 2×10^{-3}, and 7×10^{-3}, respectively.

For $n = 6$, a clear energy peak is seen at $|\sigma_{nlm}| = 8 \times 10^{-2}$ for $m = 0$, at 6×10^{-2} for $m = 2$, and at 2×10^{-2} for $m = 4$. A red shift appears in the energy peaks for the higher vertical modes. The spectral slopes in the low-frequency range are less steep than the 3 power law in $m = 0$ and 2, but the slope is steeper than 3 in $m = 4$. The slopes in the high-frequency range are about $-5/3$ for $m = 0$, and these are less steep for $m = 2$ and 4. The spectral characteristics of $n = 4$ and 2 are very similar to those for $n = 6$. The energy peaks in $n = 4$ are at 6×10^{-2} for $m = 0$, at 6×10^{-2} for $m = 2$, and at 1×10^{-2} for $m = 4$. Those in $n = 2$ are at 6×10^{-2} for $m = 0$, at 2×10^{-2} for $m = 2$, and at 5×10^{-3} for $m = 4$.

Evidently, the energy peaks are not merely associated with the largest-scale Rossby and gravity modes, but the peaks occur at the intermediate meridional scale of the Rossby modes. Since the energy flux analysis in the spectral domain provides no appropriate explanation of the peaks, it is reasonable to expect certain energy sources that correspond to those distinct energy peaks. We note that the characteristic bimodal energy peaks found in Fig. 2 and the distinct energy peaks in Fig. 4 become clearer in the transient components of atmospheric motions, E_T (not illustrated). Conversely, the spectral distributions of the quasi-stationary components, $E_S + E_A$, are rather complicated by several distinct energy peaks in the 3-D spectral domain. Therefore, the spectral peaks are more associated with the transient motions of the atmosphere.

4. Baroclinic instability on a sphere

a. Growth rate and phase speed

The spectral peaks shown in the vertical wavenumber domain and in the eigenfrequency domain in section 3 can be closely related to atmospheric baroclinic instability on a sphere. At least for synoptic waves, the role of baroclinic instability in providing the energy for disturbances, thereby creating the spectral peaks (at lease in v-component), seems to have been established (see Tanaka 1985). Figure 5 illustrates the growth rates and phase speeds analyzed from a realistic zonal basic state of the monthly mean for January 1979 (after Tanaka and Kung 1989), according to the eigenvalue problem discussed in section 2. There is a dominant unstable mode in synoptic waves with a maximum growth rate of 0.4 (day^{-1}) (e-folding time is approximately 2.5 day) and a phase speed of 10 (deg day^{-1}). This unstable mode is identified as a shallow Charney mode, M_C, on a sphere. It is reasonable to consider that this unstable mode is responsible for the energy peaks in the spectral domain of synoptic waves as shown in section 3.

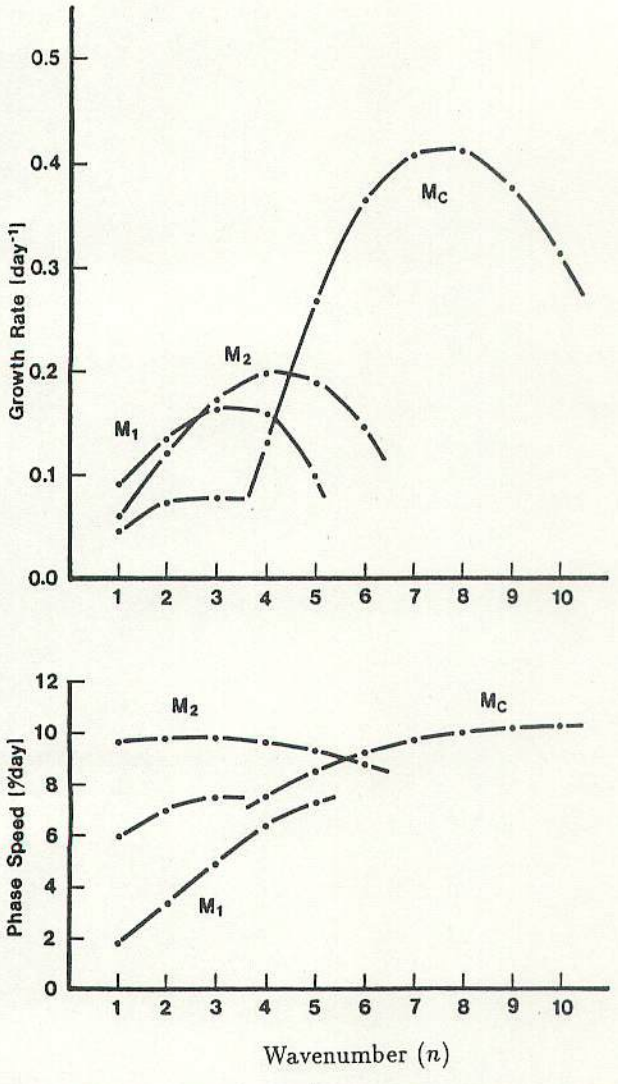

FIG. 5. Growth rates and phase speeds of the unstable modes with a monthly mean zonal basic state for January 1979 (after Tanaka and Kung 1989). The unstable modes are labeled M_C for shallow Charney modes, M_2 for dipole Charney modes, and M_1 for slow Charney modes.

Compared with the growth rates by Gall (1976) and Simmons and Hoskins (1976), present results indicate smaller growth rates beyond $n = 9$ (see Kasahara and Tanaka 1989, for $n > 10$). One of the possible reasons for this disagreement may come from the different lower boundary conditions. Most of the previous studies adopt a free-slip lower boundary condition in which an unstable mode, having its amplitude maximum at the ground, can grow quickly. In contrast, our primitive equation model adopts a nonslip lower boundary condition (7) in which the external unstable mode typical at the synoptic to short waves must damp. Since our major concern is planetary waves, whose scale is larger than the major kinetic energy source at synoptic waves, separate study should be conducted on this subject, using higher vertical resolutions and a realistic static stability.

Hartmann (1979) found different types of unstable modes M_1 and M_2, dominating planetary waves, where the shallow Charney mode approaches the neutral point and the Green mode (Green 1960) indicates small growth rates. Hartmann's results were based on an initial value problem of a quasi-geostrophic model. Tanaka and Kung (1989) confirmed his results based on an eigenvalue problem of a primitive equation model. They identified the most unstable modes at $n = 3$–4, M_2, as a dipole Charney mode (a Charney mode, but with a meridional dipole structure as in Fig. 12). In this study, the most unstable mode at $n = 1$–2, M_1, may be referred to as a slow Charney mode due to the characteristics of the slow phase speed.

The results by Tanaka and Kung (1989) were based on a spherical domain extending from the ground to about 1 mb so that the planetary waves are better resolved. Instead, they used a truncation with the Rossby mode basis alone, without the gravity mode basis, except for the Kelvin mode. This study relaxes the truncation so that the linear system includes Rossby modes of $l_R = 0$–25, westward gravity modes of $l_W = 0$–11, and eastward gravity modes of $l_E = 0$–11. This truncation is identical to the meridional resolution used for the normal mode energetics of the general circulation during the FGGE (Tanaka and Kung 1988). Accordingly, a direct comparison is possible between the theoretically expected energy peaks and the actual energy peaks presented in section 3 of this study.

In this study we examined the stability of two different symmetric zonal basic states: a zonal wind profile with a 30°-jet described by Simmons and Hoskins (1976) and an observed zonal wind profile for the monthly mean of January 1979. The 30°-jet profile has a separable structure in both the vertical and in the meridional, and is deliberately chosen to be barotropically stable. The January mean wind profile is obtained from the FGGE III-b data assimilated by the GFDL. The northern wind profile of the basic state is extended to the Southern Hemisphere so that the basic state becomes symmetric about the equator. The mean geopotential field is diagnostically evaluated so as to satisfy the geostrophic wind balance. For such symmetric zonal basic states, the eigenspace becomes decoupled into symmetric and antisymmetric solutions. The antisymmetric solution includes the antisymmetric zonal wind, antisymmetric geopotential, and symmetric meridional wind; the symmetric solution contains the opposite. We examined the antisymmetric solutions because the atmospheric eddy energy contain more antisymmetric components than symmetric components (Tanaka et al. 1986). The antisymmetric components are the even meridional indices for Rossby modes and are odd meridional indices for gravity modes. In the vertical we used the vertical indices of $m = 0$–6, according to the scheme by Tanaka and Kung (1989). The convergence of the solutions in planetary waves has been demonstrated, using this vertical res-

olution, although more vertical resolutions may be required for higher zonal wavenumbers. Table 1 lists the correspondence between the vertical indices and vertical wavenumbers, which are slightly different from those given in section 3.

The results of several growth rates and phase speeds for the 30°-jet and the January-mean basic states are listed in Table 2 for $n = 6$, 4, and 2. The results for $n = 6$ show a dominant unstable mode of the shallow Charney mode as expected from Fig. 5 and previous studies (e.g. Simmons and Hoskins 1976; Young and Houben 1989). The growth rates are 0.41 and 0.33 (day^{-1}) for the 30°-jet and the January mean basic states, respectively. The phase speed is about 8 (deg day^{-1}) for both cases. The table also lists the second unstable mode, which can be identified as a dipole Charney mode.

The results for $n = 4$ are more complicated than those of $n = 6$. The first three unstable modes with the 30°-jet basic state are identified as a shallow Charney mode, dipole Charney mode, and slow Charney mode. The results from the January-mean basic state exhibit a new type of the most unstable mode. It propagates eastward with a speed of 45° day^{-1} (i.e., 2-day period). This mode is not seen in the results from the 30°-jet basic state where barotropic instability has been ruled out. Therefore, the mode is induced by a certain feature of the realistic January basic state with a subtropical jet and a polar night jet. Hartmann (1983) investigated the barotropic instability of the polar night jet in the winter stratosphere. He found barotropically unstable modes in planetary waves with significant growth rates and a period of a few days. His barotropically unstable modes have nearly a constant phase with height and a rapidly increasing amplitude between 50 and 4 mb. The structure explains the energy maximum at the first internal vertical index, $m = 1$, as shown in Figs. 8 and 10. The structure in this study is shown in Fig. 12b for comparison. It is similar to the results by Hartmann. Therefore, this fast-moving mode can be identified as a barotropically unstable mode on a sphere. The second and third fastest-growing modes are identified as a slow Charney mode and a shallow Charney mode, respectively. There are two other fast-moving modes in the forth and fifth modes, and the sixth mode is identified as a dipole Charney mode.

The results with the January basic state for $n = 2$ similarly indicate the barotropically unstable mode as the fastest-growing mode. The second fastest growing mode shows a phase speed of 261° day^{-1} (i.e., a period of less than one day). This very fast-moving mode is identified as an unstable gravity mode, as will be clear from the prevailing gravity mode energy. The third and forth modes are identified as a slow Charney mode and a dipole Charney mode, respectively.

From the results of the growth rates and phase speeds, we found that including the gravity mode basis enhances the growth rate of the barotropically unstable mode in planetary waves. Ageostrophic components may be important, to some extent, for this mode. The results suggest that the growth rate can be underestimated in the quasi-geostrophic theory. The growth rate becomes comparable to that of low-frequency planetary waves. It is noted that even for a geostrophically balanced zonal basic state, high-frequency unstable gravity modes appear for the realistic January-mean basic state.

TABLE 2. Growth rates (day^{-1}) and phase speeds (deg day^{-1}) of the most unstable modes for the 30°-jet basic state and for the January-mean basic state. The unstable modes are labeled M_C for shallow Charney modes, M_2 for dipole Charney modes, M_1 for slow Charney modes, B for barotropically unstable modes and G for unstable gravity modes.

30° jet basic state			January-mean basic state		
Mode	Growth rate	Phase speed	Mode	Growth rate	Phase speed
	wavenumber 6				
1 M_C	0.41	7.6	1 M_C	0.33	7.7
2 M_2	0.24	7.7	2 M_2	0.17	8.3
	wavenumber 4				
1 M_C	0.21	7.8	1 B	0.20	45.1
2 M_2	0.16	6.7	2 M_1	0.18	2.7
3 M_1	0.12	7.5	3 M_C	0.17	8.2
			4 B	0.16	22.6
			5 B	0.13	33.5
			6 M_2	0.11	7.3
	wavenumber 2				
1 M_C	0.11	7.9	1 B	0.23	43.8
2 M_2	0.08	7.7	2 G	0.15	261.0
3 M_1	0.06	5.5	3 M_1	0.09	1.4
			4 M_2	0.08	8.2

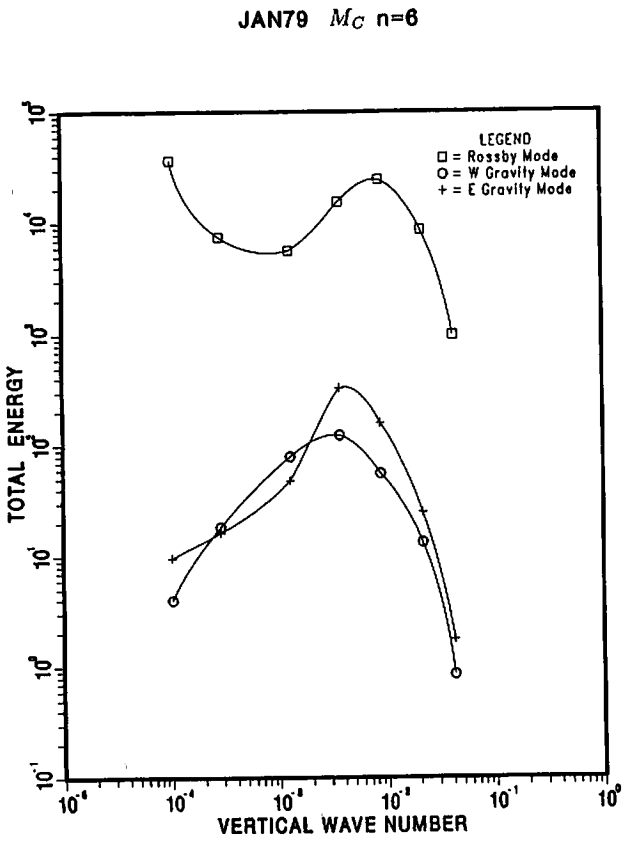

FIG. 6. Energy spectrum in the vertical wavenumber domain for $n = 6$ as in Fig. 2, but for shallow Charney mode.

High-frequency unstable modes are abundant in the unbalanced basic state: for example, the climatological mean wind and climatological mean geopotential. Since the scientific significance of unstable gravity modes is still an open question, attention is concentrated on slow, unstable modes in this study. The studies of slow, unstable modes of the primitive equations have long been hampered by the existence of high-frequency unstable gravity modes. The present 3-D spectral primitive equation model is superior for dealing with or excluding the high-frequency subspace of atmospheric motions.

b. *Energy and energy-source spectra for wavenumber 6*

The energy spectrum in the vertical wavenumber domain for the most unstable mode at $n = 6$ is illustrated in Fig. 6, using the same format as presented in Fig. 2. The spectrum indicates bimodal peaks at $h_m^{-1} = 1 \times 10^{-4}$ and 9×10^{-3}. There is an evident energy gap at $h_m^{-1} = 1 \times 10^{-3}$. The energy levels drop rapidly for the larger vertical wavenumbers. Spectral characteristics coincide with the observations shown in Fig. 2 for $n = 6$. Energy peaks in gravity modes are seen at $h_m^{-1} = 4 \times 10^{-3}$, also coinciding with observations. Hence, it is reasonable to interpret the observed bimodality in the vertical energy spectrum for $n = 6$ as the result of atmospheric baroclinic instability of the shallow Charney mode.

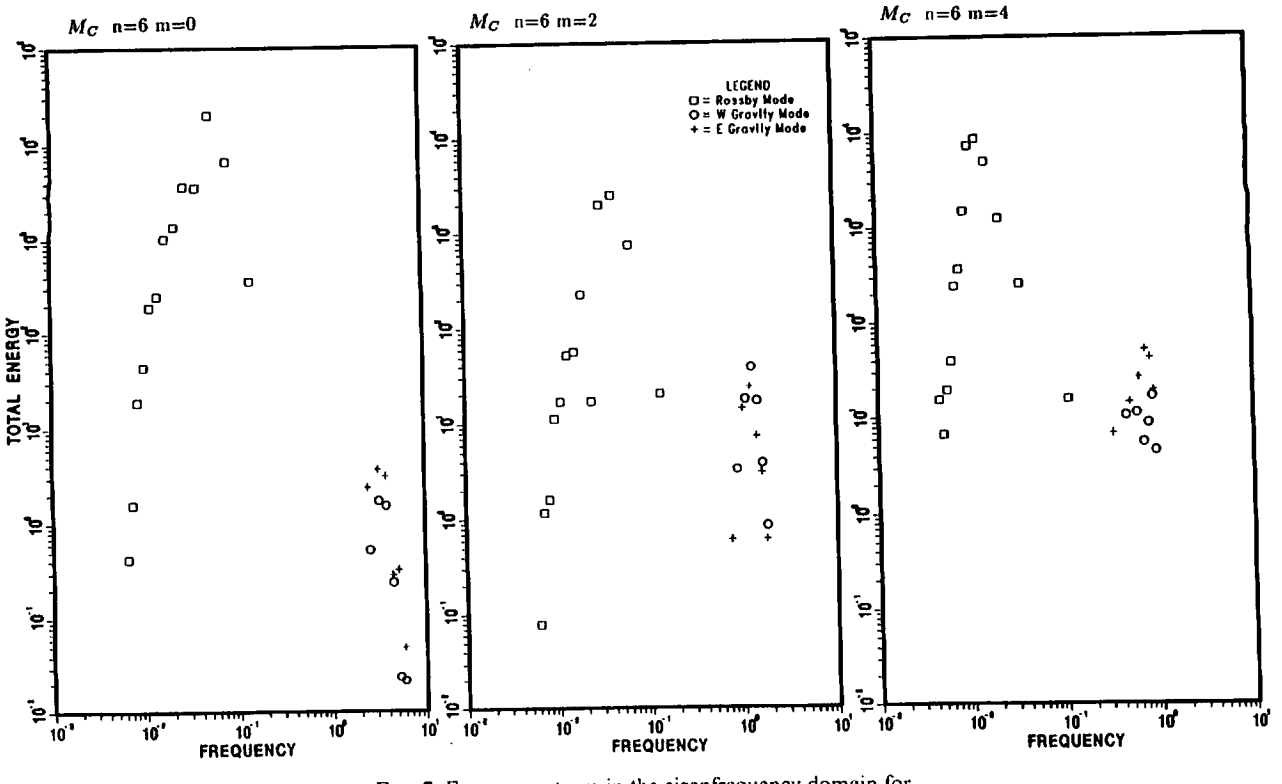

FIG. 7. Energy spectrum in the eigenfrequency domain for $n = 6$ as in Fig. 3, but for shallow Charney mode.

Figure 7 illustrates the energy spectrum in the eigenfrequency domain for the most unstable mode at $n = 6$ (compare with observations in Fig. 4). The unstable mode exhibits evident spectral peaks at $|\sigma_{nlm}| = 6 \times 10^{-2}$ for $m = 0$, at 4×10^{-2} for $m = 2$, and at 1×10^{-2} for $m = 4$. These spectral peaks coincide fairly well with the observations shown in Fig. 4. Secondary energy peaks appear in the high-frequency range. The characteristic red shift of the major peaks for the larger vertical wavenumbers is detected as it was in the observed energy spectrum. Considering the fact that the energy source has a spectral shape identical to the energy spectrum, we confirm that the observed energy peaks in the frequency domain are produced by the energy source resulting from atmospheric baroclinic instability. The present linear model, however, cannot explain the observed power law in Fig. 4 due to the lack of nonlinear interactions. In the real atmosphere, the supplied energy in the source range would cascade down to the rest of the frequency domain by nonlinear wave–wave interactions.

c. *Energy and energy-source spectra for wavenumber 4*

In conjunction with the characteristic energy peaks of $n = 6$ resulting from atmospheric baroclinic instability, the energy peaks of $n = 4$ and $n = 2$ in the vertical wavenumber and eigenfrequency domain are intriguing because the spectral features are very similar to those for $n = 6$. We have examined the energy spectra of the unstable modes for $n = 4$ listed in Table 2.

Figure 8 compares the energy spectra in the vertical wavenumber domain for the first three unstable modes; i.e., the barotropically unstable mode, slow Charney mode, and shallow Charney mode. The barotropically unstable mode exhibits an energy peak at $m = 1$. This indicates that the model structure has an energy concentration near the top of the atmosphere. The growth rate and phase speed are very similar to the results by Hartmann (1983). The first mode can be identified as a barotropically unstable mode as previously discussed. However, the energy spectrum with a peak at $h_m^{-1} = 3 \times 10^{-4}$ is different from the observed results in Fig. 2. The high-frequency nature may result in this disagreement. The second, M_1, and third, M_C, modes show bimodal energy peaks at $h_m^{-1} = 1 \times 10^{-4}$ and 9×10^{-3}. This bimodality is very similar to the results of $n = 6$ and is also consistent with observed features.

The energy spectra in the eigenfrequency domain are illustrated in Fig. 9 for the barotropically unstable mode and for the slow Charney mode at $n = 4$. The results for the barotropically unstable mode indicate distinct energy peaks at $|\sigma_{nlm}| = 4 \times 10^{-2}$ for $m = 0$, at 2×10^{-2} for $m = 2$, and at 5×10^{-3} for $m = 4$. The energy spectrum for $m = 1$ (not shown) is similar to that for $m = 0$. These eigenfrequencies are substantially different from the frequency of the unstable modes in Table 2. The high energy levels of gravity modes are notable characteristics, especially for $m = 2$. The growth rate of this mode was found to decrease with the exclusion of the gravity mode basis and the removal of the symmetric assumption for the zonal basic state.

For the slow Charney mode at $n = 4$, the distributions indicate sharp energy peaks at $|\sigma_{nlm}| = 6 \times 10^{-2}$ for $m = 20$, at 2×10^{2} for $m = 2$, and at 5×10^{-3} for $m = 4$. These energy peaks coincide with the observed values in Fig. 4. The spectral features of the slow Charney mode are very similar to the shallow Charney mode in Fig. 7.

d. *Energy and energy-source spectra for wavenumber 2*

Finally, we compared the energy spectra of the unstable modes for $n = 2$ listed in Table 2. Figure 10 illustrates the energy spectra in the vertical wavenumber domain for the first four unstable modes: the barotropically unstable mode, unstable gravity mode, slow Charney mode, and the dipole Charney mode. As in the case of $n = 4$, the barotropically unstable mode exhibits an energy peak at $m = 1$. This mode can be important in characterizing the temporal variations of planetary waves in the middle atmosphere, although the spectral distribution disagrees with observed long-term mean. This mode is not present in the 30°-jet basic state that is barotropically stable.

The second unstable mode contains the largest amount of energy in eastward propagating gravity modes. The energy levels are two orders of magnitude greater than Rossby mode energy levels. Evidently, this is an unstable gravity mode. Because the energy peak is found at $m = 1$, this unstable gravity mode is excited near the model's top (see Fig. 12c). The spectral characteristics are distinctly different from observations.

The third and forth modes are identified as a slow Charney mode and dipole Charney mode, respectively. These two unstable modes show similar energy spectra, indicating the bimodal peaks as previously seen. The bimodality appears to be a common feature in all Charney-type baroclinic instability. The energy levels of the gravity modes are negligible, which justifies the use of the quasi-geostrophic theory for these two modes. The energy levels of eastward propagating gravity modes are consistently higher than those of westward propagating gravity modes, as are observed.

The energy spectra in the eigenfrequency domain are illustrated in Fig. 11 for the slow Charney mode and the dipole Charney mode. These two energy spectra exhibit essentially the same energy peaks in the eigenfrequency domain, and both are similar to observations. The main difference between the dipole Charney mode and the slow Charney mode is the appearance of many energy gaps in the barotropic component of the dipole Charney mode. High and low energy levels appear alternately for M_2.

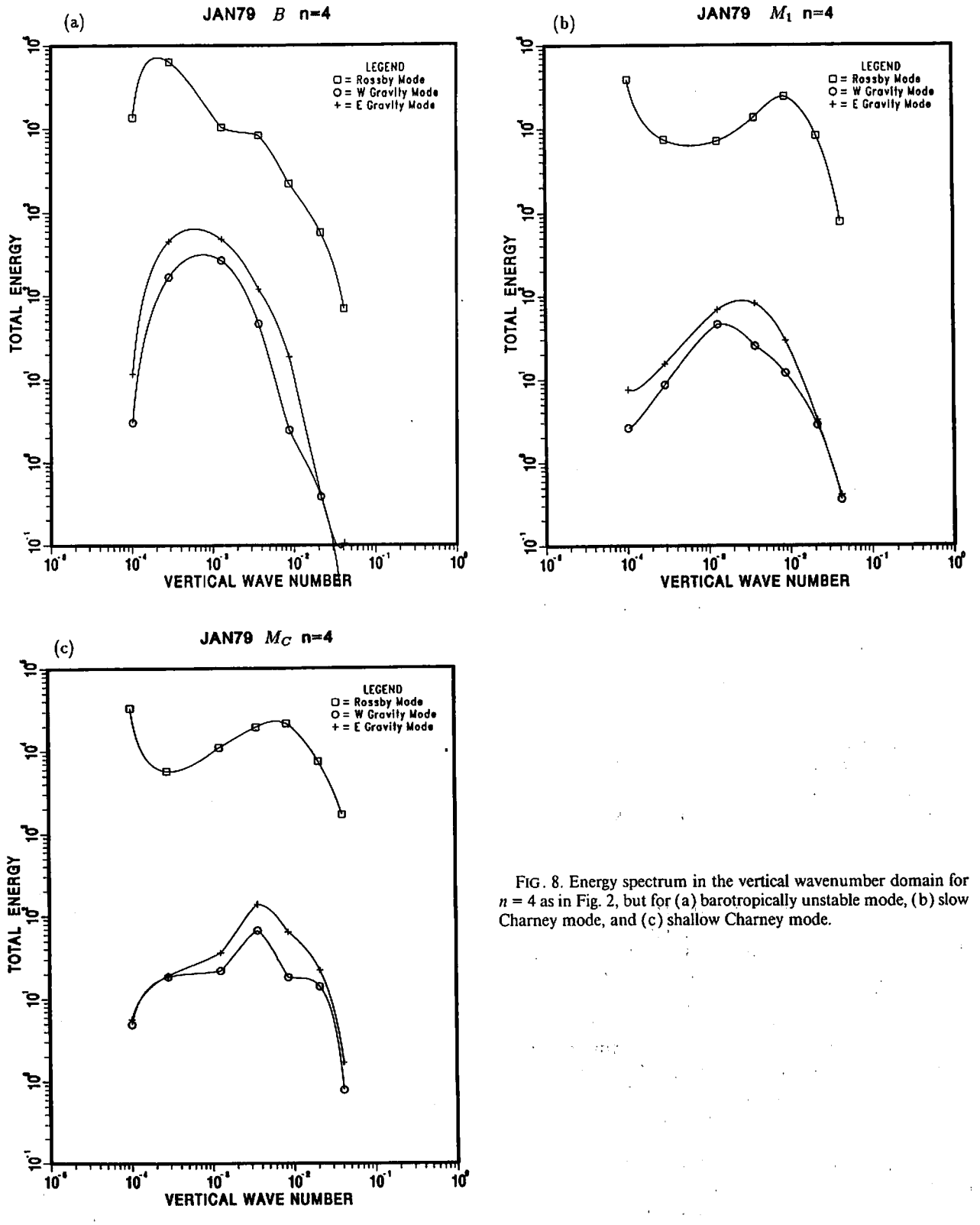

FIG. 8. Energy spectrum in the vertical wavenumber domain for $n = 4$ as in Fig. 2, but for (a) barotropically unstable mode, (b) slow Charney mode, and (c) shallow Charney mode.

e. Meridional-height structures

In order to understand the vertical and meridional energy spectra of the unstable modes, meridional-height cross-sections of the geopotential field are given in Fig. 12 for M_C of $n = 6$, and B, G, M_2 of $n = 2$. The shallow Charney mode M_C has an amplitude maximum in the troposphere, reflecting the bimodal energy peaks in the vertical spectrum. Compared with our previous studies, it is found that the inclusion of the gravity mode basis does not alter the basic structure.

The barotropically unstable mode B exhibits its am-

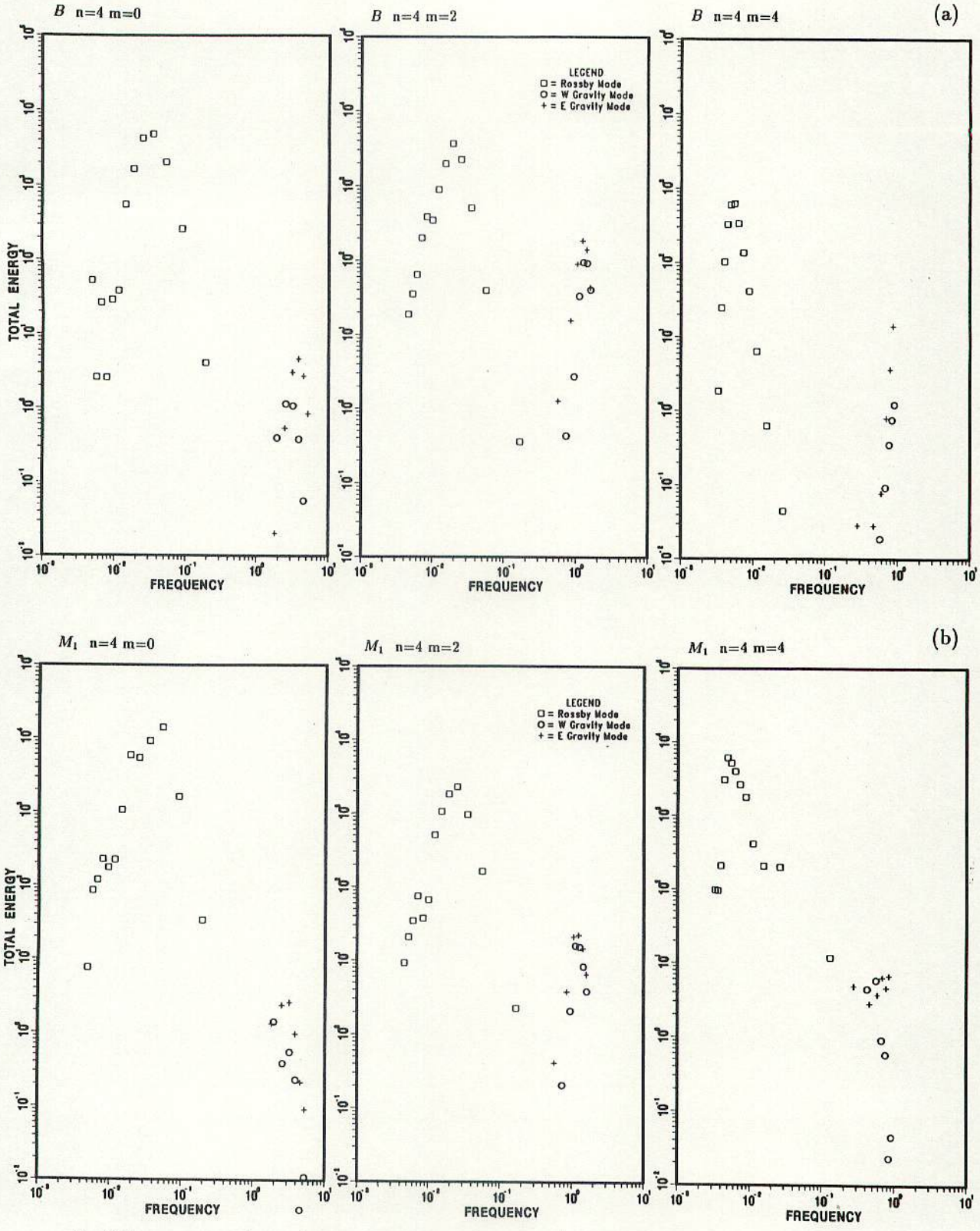

FIG. 9. Energy spectrum in the eigenfrequency domain for $n = 4$ as in Fig. 3, but for (a) barotropically unstable mode and (b) slow Charney mode.

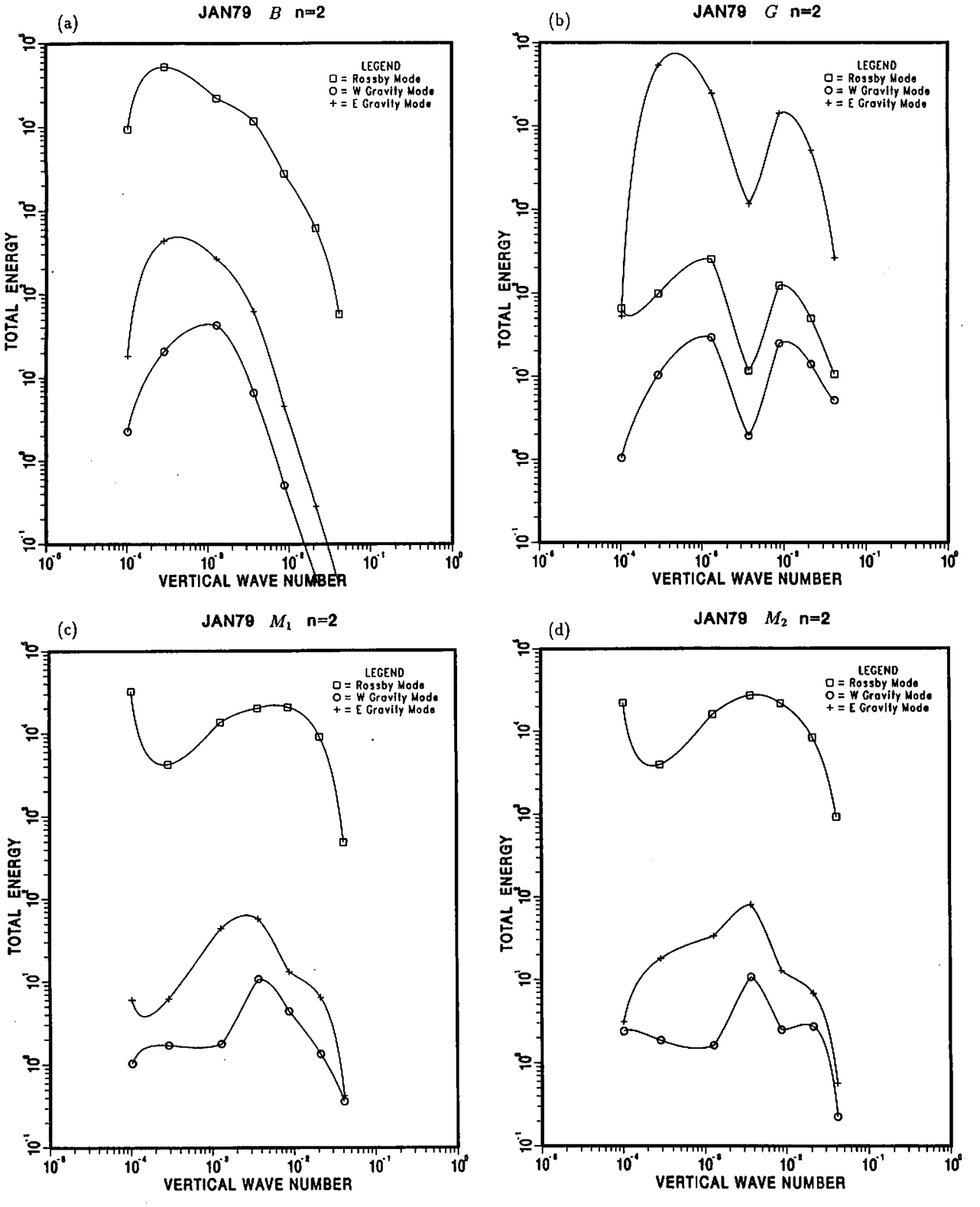

FIG. 10. Energy spectrum in the vertical wavenumber domain for $n = 2$ as in Fig. 2, but for (a) barotropically unstable mode, (b) unstable gravity mode, (c) slow Charney mode and (d) dipole Charney mode.

plitude maximum at the model's top. The structure agrees with the energy peak at $m = 1$ in Fig. 10a. It is also similar to the results by Hartmann (1983). Since a stratospheric jet core of the basic state is located near 55°N, the phase structure with a northeast–southwest tilt results in a barotropic conversion from zonal to eddy motions.

The unstable gravity mode G shows its amplitude maximum at the model's top in the low latitudes. There is no evident phase variation. Since this mode is not

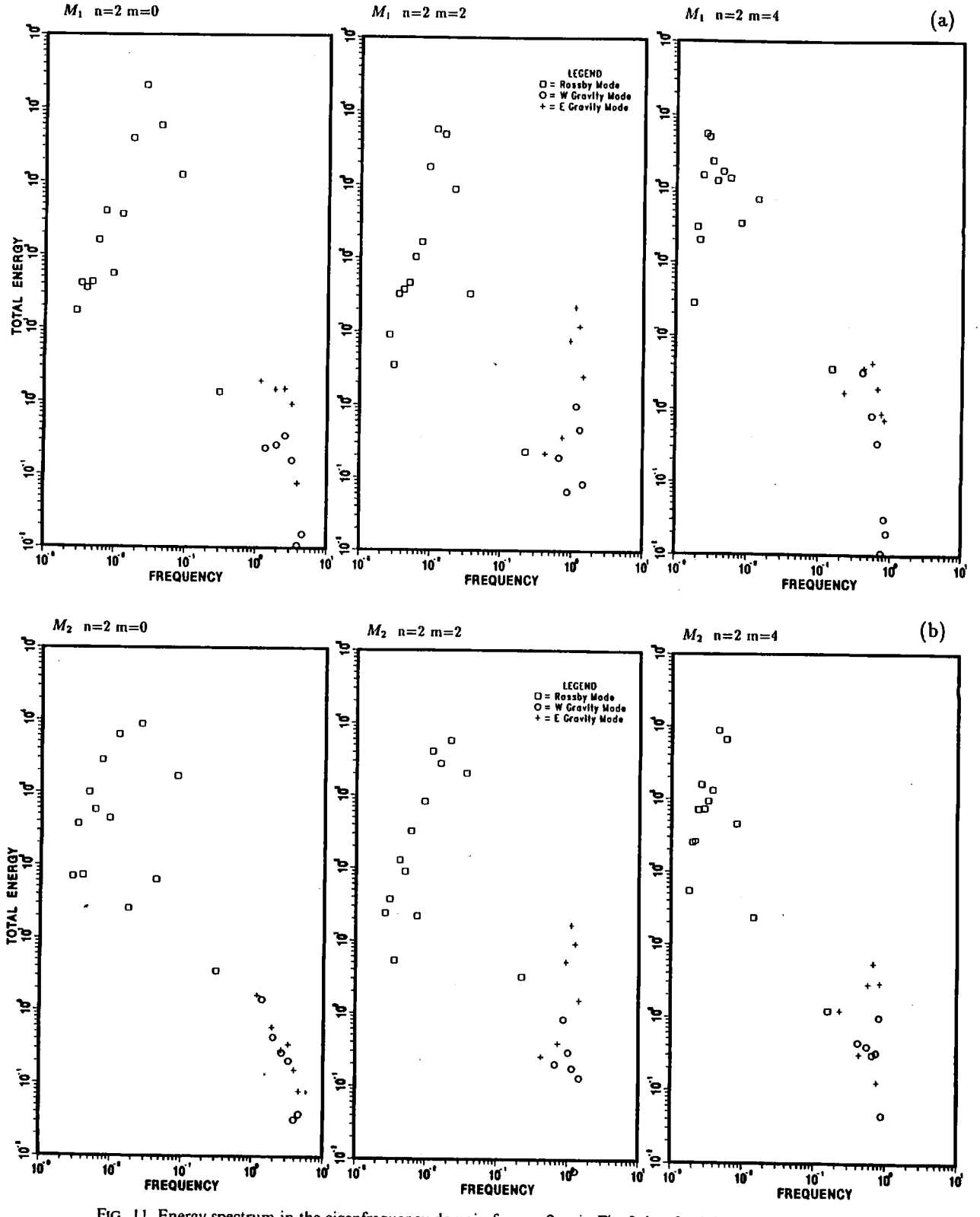

FIG. 11. Energy spectrum in the eigenfrequency domain for $n = 2$ as in Fig. 3, but for (a) slow Charney mode and (b) dipole Charney mode.

seen for the 30°-jet basic state, the mode is excited by the properties of the symmetric January basic state. The constraint of the symmetric extension of the January basic state may have caused an artificial excitation of nonphysical gravity modes. Further examination is desirable for the unstable gravity modes.

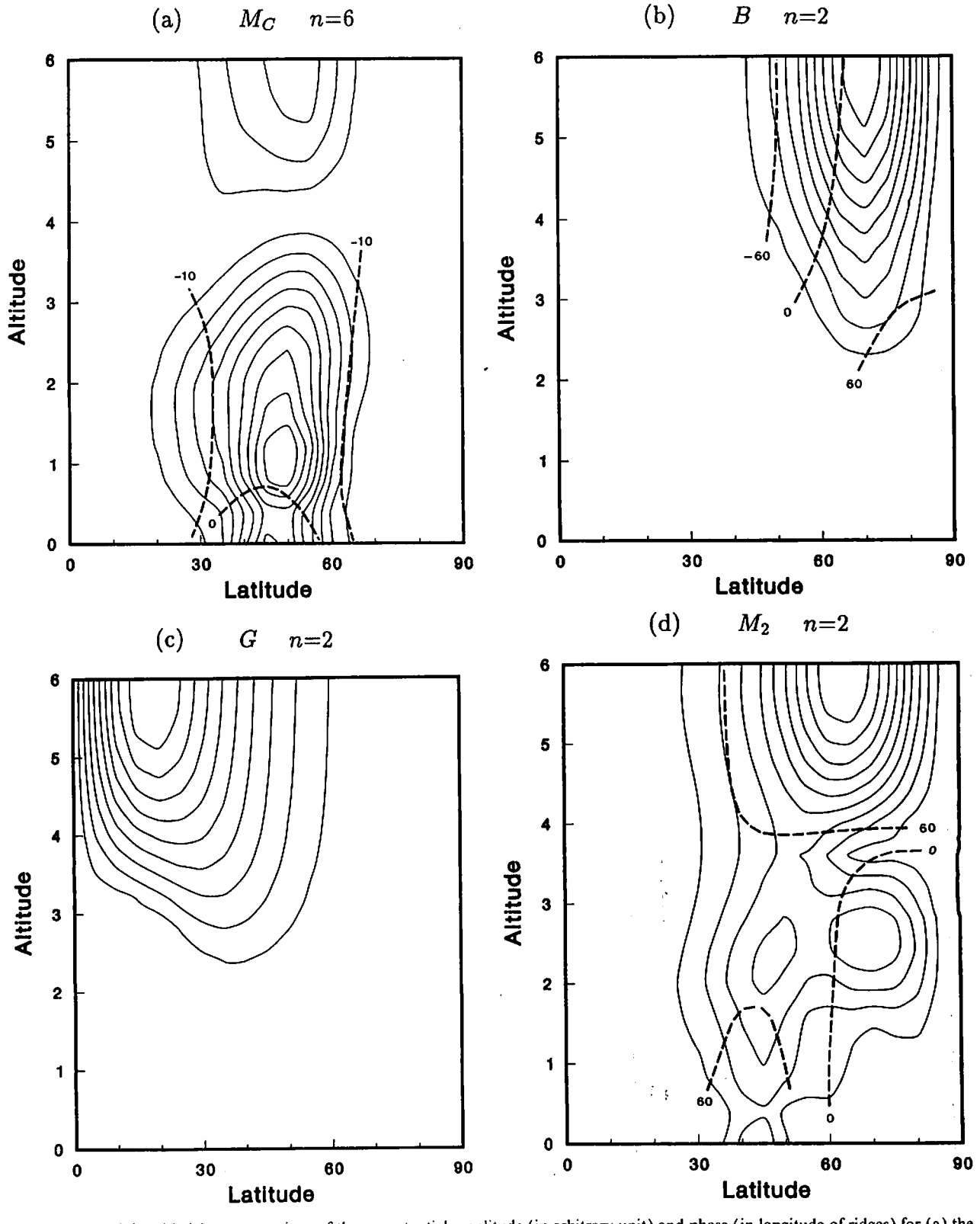

FIG. 12. Meridional-height cross sections of the geopotential amplitude (in arbitrary unit) and phase (in longitude of ridges) for (a) the shallow Charney mode M_C of wavenumber $n = 6$, (b) barotropically unstable mode B of $n = 2$, (c) unstable gravity mode G of $n = 2$, and (d) the dipole Charney mode M_2 of $n = 2$. The altitude is in $-ln\sigma$.

Finally, the dipole Charney mode M_2 exhibits a dipole amplitude configuration near the tropopause level. The actual maximum is seen at the model's top due to the density stratification effect. The structure is similar to that analyzed by Tanaka and Kung (1989) for the study of dipole blockings in the atmosphere.

5. Discussion

One of the main objectives in this study is to show the similarity between the theoretical and observational energy spectral peaks for planetary waves. This similarity implies that an orthogonal projection of the real atmosphere onto the unstable modes accounts for a large fraction of the total energy. Because the energy levels are sufficiently high in planetary waves, we can expect considerable amounts of energy supply due to the process of baroclinic instability, even through the growth rates are small. The energy conversion from quasi-stationary forcing, such as the topography and land–sea thermal contract, would be the most efficient if the unstable modes were quasi-stationary. For example, within the linear framework, the unstable mode M_1 would obtain 0.54 W m^{-2} of the energy supply from the zonal field if the observed wavenumber 2 (energy level 2.7×10^5 J m^{-2}) has the same structure as the unstable mode. This estimate is comparable to the observed baroclinic conversion (0.40 W m^{-2}) of $n = 2$ (see Kung and Tanaka 1983). The largest part of the energy supply goes to baroclinic components according to the results of the vertical energy spectrum that is proportional to the energy source spectrum.

In this regard, the physical meaning of the linearization of (22) must be reconsidered. In the real atmosphere, the single wave energy for $n = 1$–6 is about 2% to 4% of the zonal energy. The wave-mean interactions appears to dominate the wave–wave interactions in the long-term average. There are two cases when the linearization from (22) to (23) is justified: (i) when we are interested in the early stage of the full nonlinear interactions, or (ii) when the basic state is in equilibrium, and the equilibrium is maintained by well-developed wave-mean interactions. The latter case is what we require in the present study. The idea is based on the weak nonlinear resonant wave–wave interaction approximation (see Müller et al. 1986). In this approximation, it is assumed that a specific triad interaction dominates the rest of the wave–wave interactions. When the weak nonlinear resonant wave–wave interaction approximation is used, the unstable modes must also be approximately in equilibrium if small dissipative mechanisms are considered. Therefore, the growth rates of the unstable modes through this interaction should not be very large. In fact, the growth rate is not very large for planetary waves.

The origin of planetary wave energy can be topographic forcing or nonlinear upscale energy cascades from synoptic disturbances. The important energy supply from synoptic disturbances to planetary waves was confirmed by many researchers (e.g., Hansen and Chen 1982; Kung and Baker 1986; Holopainen and Fortelius 1987; Tanaka and Kung 1988). The energy conversion in baroclinic components is characterized by the downscale cascade from the zonal field to synoptic to short disturbances. In contrast, the conversion in barotropic components is characterized by the upscale cascade from synoptic disturbances to planetary waves and to zonal motions. The former is associated with the available potential energy conversion, whereas the latter with the kinetic energy conversion. The accumulated barotropic energy of planetary waves will be converted to baroclinic energy. This process is predicted by the upward group velocity of the largest-scale planetary waves (see Tanaka 1985). When the energy of planetary waves propagates upward, indicating large heat transport and a westward phase tilt, zonal available potential energy is transformed to planetary waves. During this process, which is predominately linear, the waves seem to acquire the same structure as the low-frequency unstable planetary waves because of the maximum efficiency in drawing the zonal available potential energy. It is suggested that the second interpretation of linearization as mentioned above is possible, to some extent, in planetary waves.

6. Concluding remarks

The atmospheric energy spectrum has been analyzed using the 3-D spectral primitive equations in terms of the 3-D NMF expansion. The equivalent height h_m, represents the vertical scale of motions, and the eigenfrequency of Laplace's tidal equations, σ_{nlm}, measures the 3-D scale of motions due to the intrinsic dispersion relation of Rossby waves. The method of the 3-D NMF expansion is useful for the diagnosis of the real atmosphere and the linear and nonlinear model atmospheres. The method is not restricted to a linear model for a resting atmosphere.

First, we examined the atmospheric energy spectrum for wavenumber 6 in the vertical wavenumber and eigenfrequency domain. We found characteristic bimodal energy peaks in the vertical wavenumber domain at $h_m^{-1} = 1 \times 10^{-4}$ and 7×10^{-3} (m^{-1}). We also found a distinct energy peak in the eigenfrequency domain at $|\sigma_{nlm}| = 8 \times 10^{-2}$, which separates the 3 power law in the low-frequency range and $-5/3$ power law in the high-frequency range. The results were compared with the theoretically expected energy peaks due to atmospheric baroclinic instability on a sphere. It was confirmed that the baroclinic instability of shallow Charney modes has the expected structure; i.e., characteristic bimodal energy peaks in the vertical wavenumber domain and a distinct energy peak in the eigenfrequency domain. The proof is straightforward that the spectral energy peak simultaneously describes the energy source peak for the unstable eigensolutions. Accordingly, it is understandable for wavenumber 6 that the observed energy peaks in these spectral domains result from the atmospheric baroclinic instability on a sphere. The present linear model, however, cannot explain the observed spectral slopes of 3 and $-5/3$ power law due to the lack of nonlinear interactions. The unstable mode will grow indefinitely in a linear model atmosphere, indicating an exponential increase of the energy supply at the energy peak range. In the real

atmosphere, however, the supplied energy would cascade down from the source range toward the rest of the spectral domain through the nonlinear scattering. Observed kinetic energy transformations from the synoptic waves to both the planetary and short waves (Saltzman 1970; Chen and Wiin-Nielsen 1978) support this interpretation.

Contrasted with the reasonable interpretation of energy peaks for wavenumber 6, the interpretation of the energy peaks in planetary waves has been less clear in previous research. Topographic forcing, land–sea thermal contrast, vertical propagation, barotropic and baroclinic instabilities and nonlinear wave–wave interactions are all thought to affect and complicate the spectral shape of planetary waves. However, the present analysis of the atmospheric energy spectra of planetary waves in the 3-D spectral domain indicates a rather simple spectral shape; i.e., bimodal energy peaks in the vertical wavenumber domain and a distinct energy peak in the eigenfrequency domain as seen in synoptic waves.

In order to explain the evident energy peaks in the eigenfrequency domain, we extended our analysis of atmospheric baroclinic instability for planetary waves, and compared to expected energy source spectrum with observed results. We found that the slow Charney modes and the dipole Charney modes in planetary waves exhibit the anticipated bimodal energy peaks in the vertical wavenumber domain and a sharp energy peak in the eigenfrequency domain. The resulting energy source due to atmospheric baroclinic instability coincides with the observed energy peaks of planetary waves in the eigenfrequency domain in a manner similar to synoptic waves. The results suggest that low-frequency unstable planetary waves contribute a substantial fraction of the energy peaks in the eigenfrequency domain.

The present results should not exclude the other types of energy supply for planetary waves, such as topographic forcing, land–sea thermal contrast, and nonlinear wave–wave interaction with synoptic waves. In fact, a number of observational analyses (e.g., Hansen and Chen 1982; Kung and Baker 1986; Holopainen and Fortelius 1987; Tanaka and Kung 1988) and theoretical studies (e.g., Gall et al. 1979; MacVean 1985; Young and Villere 1985; Young and Houben 1989) suggest the importance of the energy supply from synoptic to planetary waves. However, such nonlinear wave–wave interactions are inadequate to explain the statistical spectral shapes and the characteristic vertical structures discussed in this study. Linear steady solutions, on the other hand, successfully explain many features of the quasi-stationary planetary waves (see Matsuno 1970). This implies that the linear process dominates the nonlinear process for such quasi-stationary planetary waves. However, the fundamental existence of a stable steady solution is questionable under a realistic dissipation in the atmosphere. The time-mean vertical structure of planetary waves is not the steady solution, but a statistical average of episodic amplifications of transient planetary waves and chaos. The structures of the low-frequency unstable planetary waves are very similar to the linear steady solutions. The present study provides an alternative interpretation of the time-mean vertical structures of planetary waves as the occasional amplification and the phase lock of the low-frequency unstable planetary waves (see Tanaka and Kung 1989, for more detail).

Because low-frequency unstable planetary waves are free modes, they are more likely to be excited resonantly by additional external forcing than are neutral free Rossby waves. Additional quasi-stationary forcing for planetary waves tends to excite and capture the low-frequency unstable modes selectively from other numerous normal mode. The low-frequency nature of the slow Charney mode and the dipole Charney mode is more suitable to be excited and captured than the high-frequency unstable gravity modes and barotropically unstable modes that have comparable growth rates in planetary waves. Further study should be conducted making use of the fully nonlinear time-dependent models to quantify the relative importance of the energy supply from baroclinic instability and to examine the actual excitation of the low-frequency unstable modes by realistic forcing.

Acknowledgments. The authors are grateful to D. Fritts and D. Fitzgerald for reading and editing the original manuscript and to Y. Hayashi for his thorough, constructive reviews in the revision of the original manuscript. This research was supported by the National Science Foundation under Grant ATM-8923064. The second author was supported by SDIO/IST and managed by the Office of Naval Research under Contract N00014-90J-1271. The partial support came from internal funds of the Geophysical Institute, University of Alaska Fairbanks. The authors acknowledge S.-I. Akasofu for the support.

REFERENCES

Baer, F., 1972: An alternate scale representation of atmospheric energy spectra. *J. Atmos. Sci.*, **29**, 649–664.

——, 1981: Three-dimensional scaling and structure of atmospheric energetics. *J. Atmos. Sci.*, **38**, 52–68.

Boer, G. J., 1989: On exact and approximate energy equations in pressure coordinates. *Tellus*, **41A**, 97–108.

Charney, J. G., 1971: Geostrophic turbulence. *J. Atmos Sci.*, **28**, 1087–1095.

Chen, T.-C., and A. Wiin-Nielsen, 1978: On nonlinear cascades of atmospheric energy and enstrophy in a two-dimensional spectral index. *Tellus*, **30**, 313–322.

Eliasen, E., and B. Machenhauer, 1965: A study of the fluctuations of the atmospheric planetary flow patterns represented by spectral harmonics. *Tellus*, **17**, 220–238.

Gall, R., 1976: A comparison of linear baroclinic instability theory with the eddy statistics of a general circulation model. *J. Atmos. Sci.*, **33**, 349–373.

——, R. Blakeslee and R. C. J. Somerville, 1979: Cyclone-scale forcing of ultralong waves. *J. Atmos Sci.*, **36**, 1692–1698.

Green, J. S. A., 1960: A problem in baroclinic instability. *Quart. J. Roy. Meteor. Soc.*, **86**, 237–251.

Hansen, A. R., and T.-C. Chen, 1982: A spectral energetics analysis of atmospheric blocking. *Mon. Wea Rev.*, **110**, 1146–1165.
Hartmann, D. L., 1979: Baroclinic instability of realistic zonal-mean state to planetary waves. *J. Atmos. Sci.*, **36**, 2336–2349.
——, 1983: Barotropic instability of the polar night jet. *J. Atmos. Sci.*, **40**, 817–835.
Hayashi, Y., 1980: Estimation of nonlinear energy transfer spectra by the cross spectral method. *J. Atmos. Sci.*, **37**, 299–307.
Herring, J. R., 1980: Statistical theory of quasi-geostrophic turbulence. *J. Atmos. Sci.*, **37**, 969–977.
Holmström, I., 1963: On a method for parametric representation of the state of the atmosphere. *Tellus*, **15**, 127–149.
Holopainen, E., and C. Fortelius, 1987: High-frequency transient eddies and blocking. *J. Atmos. Sci.*, **44**, 1623–1645.
Holton, J. R., 1975:—The dynamic meteorology of the stratosphere and mesosphere, *Meteor. Monogr.*, 37, Amer. Meteor. Soc., 218 pp.
Hua, B. L., and D. B. Haidvogel, 1986: Numerical simulation of the vertical structure of quasi-geostrophic turbulence. *J. Atmos Sci.*, **43**, 2923–2936.
Kasahara, A., 1976: Normal modes of ultralong waves in the atmosphere. *Mon. Wea. Rev.*, **104**, 669–690.
——, 1978: Further studies on a spectral model of the global barotropic primitive equations with Hough harmonic expansions. *J. Atmos. Sci.*, **35**, 2043–2051.
——, 1984: The linear response of a stratified global atmosphere to tropical thermal forcing. *J. Atmos. Sci.*, **41**, 2217–2237.
——, and K. Puri, 1981: Spectral representation of three-dimensional global data by expansion in normal mode functions. *Mon. Wea. Rev.*, **109**, 37–51.
——, and H. L. Tanaka, 1989: Application of vertical normal mode expansion to problems of baroclinic instability. *J. Atmos. Sci.*, **46**, 489–510.
Kraichnan, R. H., 1967: Inertial range in two-dimensional turbulence. *Phys. Fluids*, **10**, 1417–1423.
——, and D. Montgomery, 1980: Two-dimensional turbulence. *Rev. Prog. Phys.*, **43**, 547–619.
Kung, E. C., 1988: Spectral energetics of the general circulation and time spectra of transient waves during the FGGE year. *J. Climate*, **1**, 5–19.
——, and W. E. Baker, 1986: Spectral energetics of the observed and simulated northern hemisphere general circulation during blocking episodes. *J. Atmos. Sci.*, **43**, 2792–2812.
——, and H. Tanaka, 1983: Energetics analysis of the global circulation during the special observational period of FGGE. *J. Atmos Sci.*, **40**, 2575–2592.
——, and ——, 1984: Spectral characteristics and meridional variations of energy transformations during the first and second special observation periods of FGGE. *J. Atmos. Sci.*, **41**, 1836–1849.
Leith, C. E., 1968: Diffusion approximation for two-dimensional turbulence. *Phys. Fluids*, **11**, 671–672.
Lilly, D. K., 1971: Numerical simulation of developing and decaying two-dimensional turbulence *J. Fluid Mech.*, **45**, 2, 395–415.
——, 1989: Two-dimensional turbulence generated by energy sources at two scales. *J. Atmos. Sci.*, **46**, 2026–2030.
Lindzen, R. S., B. Farrel and K. K. Tung, 1980: The concept of wave overreflection and its application to baroclinic instability. *J. Atmos. Sci.*, **37**, 44–63.
MacVean, M. K., 1985: Long-wave growth by baroclinic processes. *J. Atmos. Sci.*, **42**, 1089–1101.
Matsuno, T., 1970: Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *J. Atmos. Sci.*, **27**, 871–883.
Müller, P., G. Holloway, F. Henyey and N. Pomphrey, 1986: Nonlinear interaction among internal gravity waves. *Rev. Geophys.*, **24**, 495–536.
Saltzman, B., 1957: Equations governing the energetics of the larger scales of atmospheric turbulence in the domain of wave number. *J. Meteor.*, **14**, 513–523.
——, 1970: Large-scale atmospheric energetics in the wavenumber domain. *Rev. Geophys. Space Phys.*, **8**, 289–302.
Silva Dias, P. L., and J. P. Bonatti, 1986: Vertical mode decomposition and model resolution. *Tellus*, **38A**, 205–214.
Simmons, A. J., and B. J. Hoskins, 1976: Baroclinic instability on the sphere: Normal modes of the primitive and quasi-geostrophic equations. *J. Atmos. Sci.*, **33**, 1454–1477.
Staniforth, A., M. Béland and J. Côté, 1985: An analysis of the vertical structure equation in sigma coordinates. *Atmosphere: Atmos.-Ocean*, **23**, 323–358.
Tanaka, H., 1985: Global energetics analysis by expansion into three-dimensional normal mode functions during the FGGE winter. *J. Meteorol. Soc. Jpn.*, **63**, 180–200.
——, and E. C. Kung, 1988: Normal mode energetics of the general circulation during the FGGE year. *J. Atmos. Sci.*, **45**, 3723–3736.
——, and ——, 1989: A study of low-frequency unstable planetary waves in realistic zonal and zonally varying basic states. *Tellus*, **41A**, 179–199.
——, —— and W. E. Baker, 1986: Energetics analysis of the observed and simulated general circulation using three-dimensional normal mode expansion. *Tellus*, **38A**, 412–428.
Wiin-Nielsen, A., 1967: On the annual variation and spectral distribution of atmospheric energy. *Tellus*, **19**, 540–559.
Young, R. E., and G. L. Villere, 1985: Nonlinear forcing of planetary scale waves by amplifying unstable baroclinic eddies. *J. Atmos. Sci.*, **42**, 1991–2006.
——, and H. Houben, 1989: Dynamics of planetary-scale baroclinic waves during southern hemisphere winter. *J. Atmos. Sci.*, **46**, 1365–1383.